

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:32:42 ; Search time 78.5 Seconds  
(without alignments)  
27.986 Million cell updates/sec

Title: US-10-771-242-293

Perfect score: 9

Sequence: 1 RXXXX 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :

A\_Geneseq\_21:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*  
9: geneseqp2005s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5	55.6	2	2	Adh29628 Swinepox
2	5	55.6	2	4	Aam98426 Humah pep
3	5	55.6	2	4	Aab91738 Opioild pe
4	5	55.6	2	5	Abg93548 Humah P-g
5	5	55.6	2	5	Abg93543 Humah P-g
6	5	55.6	2	5	Abg93498 Humah P-g
7	5	55.6	2	5	Abg93615 Humah P-g
8	5	55.6	2	5	Abg93621 Humah P-g
9	5	55.6	2	5	Abg93576 Humah P-g
10	5	55.6	2	5	Abg93550 Humah P-g
11	5	55.6	2	5	Abp53564 Flaamid p
12	5	55.6	2	6	Abp39485 AlphaS1 c
13	5	55.6	2	7	Aad39209 Angtogene
14	5	55.6	2	7	Adl98393 Humah leu
15	5	55.6	2	7	Adw36641 HLA bindi
16	5	55.6	2	7	Adw36964 HLA bindi
17	5	55.6	2	7	Adw37285 HLA bindi
18	5	55.6	3	1	Aap30601 Sequence
19	5	55.6	3	1	Aap90668 New antih
20	5	55.6	3	1	Aap90665 New antih
21	5	55.6	3	1	Aap97811 Sequence
22	5	55.6	3	2	Aar00718 Core repe
23	5	55.6	3	2	Aar04607 Antiviral
24	5	55.6	3	2	Aar10543 Hypotensi

25	5	55.6	3	2	AAR32271	Soybean g
26	5	55.6	3	2	AAR36707	Adhesion
27	5	55.6	3	2	AAR53144	RGD pep1
28	5	55.6	3	2	AAR38487	Human RDS
29	5	55.6	3	2	AAR42569	Peptide c
30	5	55.6	3	2	AAR30753	IGF-mast
31	5	55.6	3	2	AAR48960	NI4-3 tru
32	5	55.6	3	2	AAR46824	Phytase d
33	5	55.6	3	2	AAR44666	Platelet
34	5	55.6	3	2	AAR63264	Thrombin
35	5	55.6	3	2	AAR63263	Thrombin
36	5	55.6	3	2	AAR63261	Thrombin
37	5	55.6	3	2	AAR63260	Thrombin
38	5	55.6	3	2	AAR63262	Thrombin
39	5	55.6	3	2	AAR51440	IGF-1 ana
40	5	55.6	3	2	AAR61094	ACE-inhib
41	5	55.6	3	2	AAR61092	ACE-inhib
42	5	55.6	3	2	AAR61093	ACE-inhib
43	5	55.6	3	2	AAR73934	Novel tri
44	5	55.6	3	2	AAR57458	Lactoferr
45	5	55.6	3	2	AAR57450	Lactoferr
46	5	55.6	3	2	AAR48523	Lactoferr
47	5	55.6	3	2	AAR48527	Lactoferr
48	5	55.6	3	2	AAR48519	Lactoferr
49	5	55.6	3	2	AAR58575	Angiotens
50	5	55.6	3	2	AAR58571	Angiotens
51	5	55.6	3	2	AAR58576	Angiotens
52	5	55.6	3	2	AAR58578	Angiotens
53	5	55.6	3	2	AAR58573	Angiotens
54	5	55.6	3	2	AAR58569	Angiotens
55	5	55.6	3	2	AAR58570	Angiotens
56	5	55.6	3	2	AAR58572	Angiotens
57	5	55.6	3	2	AAR58577	Angiotens
58	5	55.6	3	2	AAR84695	Bovine la
59	5	55.6	3	2	AAR84687	Bovine la
60	5	55.6	3	2	AAR84691	Bovine la
61	5	55.6	3	2	AAR69779	Thrombos
62	5	55.6	3	2	AAR70472	Cancer me
63	5	55.6	3	2	AAR82919	Non-RGD,
64	5	55.6	3	2	AAR85069	Calcium i
65	5	55.6	3	2	AAR64657	HPP3 pept
66	5	55.6	3	2	AAR64722	HPP3 pept
67	5	55.6	3	2	AAR90597	Lactoferr
68	5	55.6	3	2	AAR90605	Lactoferr
69	5	55.6	3	2	AAR90601	Lactoferr
70	5	55.6	3	2	AAR99827	Mutant ch
71	5	55.6	3	2	AAR99828	Active do
72	5	55.6	3	2	AAR98543	Peptide f
73	5	55.6	3	2	AAR98547	Peptide f
74	5	55.6	3	2	AAR98551	Peptide f
75	5	55.6	3	2	AAR88897	Small syn
76	5	55.6	3	2	AAR88899	Small syn
77	5	55.6	3	2	AAR88900	Small syn
78	5	55.6	3	2	AAR88898	Small syn
79	5	55.6	3	2	AAW11094	Platelet-
80	5	55.6	3	2	AAW00269	Cytokine
81	5	55.6	3	2	AAW25173	RGD-pepti
82	5	55.6	3	2	AAW18871	Peptide f
83	5	55.6	3	2	AAW33241	Anaigesic
84	5	55.6	3	2	AAW31314	Human int
85	5	55.6	3	2	AAW31143	Platelet-
86	5	55.6	3	2	AAW56177	Anti-infl
87	5	55.6	3	2	AAW56181	Anti-infl
88	5	55.6	3	2	AAW56183	Anti-infl
89	5	55.6	3	2	AAW56186	Anti-infl
90	5	55.6	3	2	AAW56196	Anti-infl
91	5	55.6	3	2	AAW56201	Anti-infl
92	5	55.6	3	2	AAW56234	Anti-infl
93	5	55.6	3	2	AAW56200	Anti-infl
94	5	55.6	3	2	AAW56209	Anti-infl
95	5	55.6	3	2	AAW56233	Anti-infl
96	5	55.6	3	2	AAW56188	Anti-infl
97	5	55.6	3	2	AAW56188	Anti-infl

98	5	55.6	3	2	AAW56197	Anti-infl	171	5	55.6	3	2	AAW63167	Mouse mas
99	5	55.6	3	2	AAW56207	Anti-infl	172	5	55.6	3	2	AAW63169	Mouse mas
100	5	55.6	3	2	AAW56220	Anti-infl	173	5	55.6	3	2	AAW63161	Mouse mas
101	5	55.6	3	2	AAW56239	Anti-infl	174	5	55.6	3	2	AAW63168	Mouse mas
102	5	55.6	3	2	AAW56175	Anti-infl	175	5	55.6	3	2	AAW76946	Fusion im
103	5	55.6	3	2	AAW56190	Anti-infl	176	5	55.6	3	2	AAW71120	Peptide A
104	5	55.6	3	2	AAW56195	Anti-infl	177	5	55.6	3	2	AAW41762	Colony st
105	5	55.6	3	2	AAW56230	Anti-infl	178	5	55.6	3	2	AAW41282	Apoptosis
106	5	55.6	3	2	AAW56246	Anti-infl	179	5	55.6	3	2	AAW41286	Apoptosis
107	5	55.6	3	2	AAW56185	Anti-infl	180	5	55.6	3	2	AAW41278	Apoptosis
108	5	55.6	3	2	AAW56191	Anti-infl	181	5	55.6	3	2	ADI30304	Peptide r
109	5	55.6	3	2	AAW56193	Anti-infl	182	5	55.6	3	2	ADI30311	Peptide r
110	5	55.6	3	2	AAW56198	Anti-infl	183	5	55.6	3	2	AAI30593	Amino aci
111	5	55.6	3	2	AAW56213	Anti-infl	184	5	55.6	3	2	AAW84193	Peptide c
112	5	55.6	3	2	AAW56229	Anti-infl	185	5	55.6	3	2	AAI30549	Amino aci
113	5	55.6	3	2	AAW56240	Anti-infl	186	5	55.6	3	2	AAI30308	Peptide r
114	5	55.6	3	2	AAW56178	Anti-infl	187	5	55.6	3	2	AAI30311	Peptide r
115	5	55.6	3	2	AAW56202	Anti-infl	188	5	55.6	3	2	AAI30311	Peptide r
116	5	55.6	3	2	AAW56212	Anti-infl	189	5	55.6	3	2	AAI30311	Peptide r
117	5	55.6	3	2	AAW56216	Anti-infl	190	5	55.6	3	2	AAI30311	Peptide r
118	5	55.6	3	2	AAW56232	Anti-infl	191	5	55.6	3	2	AAI30311	Peptide r
119	5	55.6	3	2	AAW56241	Anti-infl	192	5	55.6	3	2	AAI30311	Peptide r
120	5	55.6	3	2	AAW56172	Anti-infl	193	5	55.6	3	2	AAI30311	Peptide r
121	5	55.6	3	2	AAW56173	Anti-infl	194	5	55.6	3	2	AAI30311	Peptide r
122	5	55.6	3	2	AAW56180	Anti-infl	195	5	55.6	3	2	AAI30311	Peptide r
123	5	55.6	3	2	AAW56182	Anti-infl	196	5	55.6	3	2	AAI30311	Peptide r
124	5	55.6	3	2	AAW56203	Anti-infl	197	5	55.6	3	2	AAI30311	Peptide r
125	5	55.6	3	2	AAW56204	Anti-infl	198	5	55.6	3	2	AAI30311	Peptide r
126	5	55.6	3	2	AAW56208	Anti-infl	199	5	55.6	3	2	AAI30311	Peptide r
127	5	55.6	3	2	AAW56210	Anti-infl	200	5	55.6	3	2	AAI30311	Peptide r
128	5	55.6	3	2	AAW56221	Anti-infl	201	5	55.6	3	2	AAI30311	Peptide r
129	5	55.6	3	2	AAW56236	Anti-infl	202	5	55.6	3	2	AAI30311	Peptide r
130	5	55.6	3	2	AAW56206	Anti-infl	203	5	55.6	3	2	AAI30311	Peptide r
131	5	55.6	3	2	AAW56214	Anti-infl	204	5	55.6	3	2	AAI30311	Peptide r
132	5	55.6	3	2	AAW56218	Anti-infl	205	5	55.6	3	2	AAI30311	Peptide r
133	5	55.6	3	2	AAW56222	Anti-infl	206	5	55.6	3	2	AAI30311	Peptide r
134	5	55.6	3	2	AAW56235	Anti-infl	207	5	55.6	3	2	AAI30311	Peptide r
135	5	55.6	3	2	AAW56238	Anti-infl	208	5	55.6	3	2	AAI30311	Peptide r
136	5	55.6	3	2	AAW56248	Anti-infl	209	5	55.6	3	2	AAI30311	Peptide r
137	5	55.6	3	2	AAW56187	Anti-infl	210	5	55.6	3	2	AAI30311	Peptide r
138	5	55.6	3	2	AAW56227	Anti-infl	211	5	55.6	3	2	AAI30311	Peptide r
139	5	55.6	3	2	AAW56194	Anti-infl	212	5	55.6	3	2	AAI30311	Peptide r
140	5	55.6	3	2	AAW56199	Anti-infl	213	5	55.6	3	2	AAI30311	Peptide r
141	5	55.6	3	2	AAW56211	Anti-infl	214	5	55.6	3	2	AAI30311	Peptide r
142	5	55.6	3	2	AAW56245	Anti-infl	215	5	55.6	3	2	AAI30311	Peptide r
143	5	55.6	3	2	AAW56189	Anti-infl	216	5	55.6	3	2	AAI30311	Peptide r
144	5	55.6	3	2	AAW56219	Anti-infl	217	5	55.6	3	2	AAI30311	Peptide r
145	5	55.6	3	2	AAW56174	Anti-infl	218	5	55.6	3	2	AAI30311	Peptide r
146	5	55.6	3	2	AAW56205	Anti-infl	219	5	55.6	3	2	AAI30311	Peptide r
147	5	55.6	3	2	AAW56217	Anti-infl	220	5	55.6	3	2	AAI30311	Peptide r
148	5	55.6	3	2	AAW56223	Anti-infl	221	5	55.6	3	2	AAI30311	Peptide r
149	5	55.6	3	2	AAW56224	Anti-infl	222	5	55.6	3	2	AAI30311	Peptide r
150	5	55.6	3	2	AAW56231	Anti-infl	223	5	55.6	3	2	AAI30311	Peptide r
151	5	55.6	3	2	AAW56244	Anti-infl	224	5	55.6	3	2	AAI30311	Peptide r
152	5	55.6	3	2	AAW56237	Anti-infl	225	5	55.6	3	2	AAI30311	Peptide r
153	5	55.6	3	2	AAW56237	Anti-infl	226	5	55.6	3	2	AAI30311	Peptide r
154	5	55.6	3	2	AAW56242	Anti-infl	227	5	55.6	3	2	AAI30311	Peptide r
155	5	55.6	3	2	AAW56243	Anti-infl	228	5	55.6	3	2	AAI30311	Peptide r
156	5	55.6	3	2	AAW56176	Anti-infl	229	5	55.6	3	2	AAI30311	Peptide r
157	5	55.6	3	2	AAW56247	Anti-infl	230	5	55.6	3	2	AAI30311	Peptide r
158	5	55.6	3	2	AAW56171	Anti-infl	231	5	55.6	3	2	AAI30311	Peptide r
159	5	55.6	3	2	AAW56184	Anti-infl	232	5	55.6	3	2	AAI30311	Peptide r
160	5	55.6	3	2	AAW56192	Anti-infl	233	5	55.6	3	2	AAI30311	Peptide r
161	5	55.6	3	2	AAW56215	Anti-infl	234	5	55.6	3	2	AAI30311	Peptide r
162	5	55.6	3	2	AAW56225	Anti-infl	235	5	55.6	3	2	AAI30311	Peptide r
163	5	55.6	3	2	AAW56226	Anti-infl	236	5	55.6	3	2	AAI30311	Peptide r
164	5	55.6	3	2	AAW56228	Anti-infl	237	5	55.6	3	2	AAI30311	Peptide r
165	5	55.6	3	2	AAW64738	Argiotens	238	5	55.6	3	2	AAI30311	Peptide r
166	5	55.6	3	2	AAW48589	Integrin	239	5	55.6	3	2	AAI30311	Peptide r
167	5	55.6	3	2	AAW52450	Loop regi	240	5	55.6	3	2	AAI30311	Peptide r
168	5	55.6	3	2	AAW52449	Loop regi	241	5	55.6	3	2	AAI30311	Peptide r
169	5	55.6	3	2	AAW52447	Loop regi	242	5	55.6	3	2	AAI30311	Peptide r
170	5	55.6	3	2	AAW61942	PPI bindi	243	5	55.6	3	2	AAI30311	Peptide r

244	5	55.6	3	5	ABG77863	Targettin	317	5	55.6	3	7	ADW37278	HLA bindi
245	5	55.6	3	5	ABG77883	Targettin	318	5	55.6	3	7	ADW36968	HLA bindi
246	5	55.6	3	5	ABG77889	Targettin	319	5	55.6	3	7	ADI36142	Human ang
247	5	55.6	3	5	ABG77792	Targettin	320	5	55.6	3	8	ADJ62029	Dut prote
248	5	55.6	3	5	ABG77864	Targettin	321	5	55.6	3	8	ADM01233	Targeted
249	5	55.6	3	5	ABG77785	Targettin	322	5	55.6	3	8	ADM01247	Targeted
250	5	55.6	3	5	ABG77807	Targettin	323	5	55.6	3	8	ADM01248	Targeted
251	5	55.6	3	5	ABG77887	Targettin	324	5	55.6	3	8	ADM01249	Targeted
252	5	55.6	3	5	ABG77890	Targettin	325	5	55.6	3	8	ADM32999	Human imm
253	5	55.6	3	5	ABG77786	Targettin	326	5	55.6	3	8	ADM33004	Human imm
254	5	55.6	3	5	ABG77779	Targettin	327	5	55.6	3	8	ADM33000	Human ggl
255	5	55.6	3	5	ABG77780	Targettin	328	5	55.6	3	8	ADN17303	Human ggl
256	5	55.6	3	5	ABG77814	Targettin	329	5	55.6	3	8	ADP30381	Human sec
257	5	55.6	3	5	ABG77845	Targettin	330	5	55.6	3	8	ADP29325	Human sec
258	5	55.6	3	5	ABG77846	Targettin	331	5	55.6	3	8	ADO34674	Epstein B
259	5	55.6	3	5	ABG77805	Targettin	332	5	55.6	3	8	ADQ59611	Peptide 1
260	5	55.6	3	5	ABG77839	Targettin	333	5	55.6	3	8	ADRO4670	Factor XA
261	5	55.6	3	5	ABG77852	Targettin	334	5	55.6	3	8	ADR04679	Factor XA
262	5	55.6	3	5	ABG77783	Targettin	335	5	55.6	3	8	ADS85334	PCAV puta
263	5	55.6	3	5	ABG77811	Targettin	336	5	55.6	3	8	ADR68276	Murine 26
264	5	55.6	3	5	ABG77488	Targettin	337	5	55.6	3	8	ADS82723	Arabidops
265	5	55.6	3	5	ABG77775	Targettin	338	5	55.6	3	8	ADU15638	MUC1 cyto
266	5	55.6	3	5	ABG77808	Targettin	339	5	55.6	3	8	ADU15617	MUC1 cyto
267	5	55.6	3	5	ABG77876	Targettin	340	5	55.6	3	8	ADU15628	MUC1 cyto
268	5	55.6	3	5	ABG77835	Targettin	341	5	55.6	3	8	ADT51230	G protein
269	5	55.6	3	5	ABG77843	Targettin	342	5	55.6	3	8	ADT51239	G protein
270	5	55.6	3	5	AAU85775	Angiotens	343	5	55.6	3	8	ADT51228	G protein
271	5	55.6	3	6	ABU72609	Novel pro	344	5	55.6	3	9	ADV52535	Gamma-sec
272	5	55.6	3	6	ABU72668	Novel pro	345	5	55.6	3	9	ADY29695	Antibacte
273	5	55.6	3	6	ABU90757	Peptide #	346	5	55.6	3	9	ADY29701	Antibacte
274	5	55.6	3	6	ABU13829	Thrombos	347	5	55.6	3	9	ADX39036	Lactate d
275	5	55.6	3	6	ABU13855	Thrombos	348	5	55.6	3	9	ADY64208	Bovine la
276	5	55.6	3	6	ABU13854	Thrombos	349	5	55.6	3	9	ADY64204	Bovine la
277	5	55.6	3	6	ABU60858	Peptide p	350	5	55.6	3	9	ADY64200	Bovine la
278	5	55.6	3	6	ABP99827	Human sec	351	5	55.6	3	9	ADY99974	Rat neuro
279	5	55.6	3	6	ABR39486	Alphas1 c	352	5	55.6	3	9	AEA21530	Human ant
280	5	55.6	3	6	ABU11422	Angiotens	353	5	55.6	3	9	AEA35151	LPS bindi
281	5	55.6	3	6	ABR55066	MMP subet	354	5	55.6	3	9	AEA39119	Flaviviri
282	5	55.6	3	6	ABR55098	MMP subet	355	5	55.6	3	9	AEA39121	Flaviviri
283	5	55.6	3	6	ABR55099	MMP subet	356	5	55.6	3	9	AEA39117	Flaviviri
284	5	55.6	3	6	ABR55060	MMP subet	357	5	55.6	3	9	AEA39123	Flaviviri
285	5	55.6	3	6	ABR55101	MMP subet	358	5	55.6	4	1	AAP10137	Sequence
286	5	55.6	3	6	ABR55068	MMP subet	359	5	55.6	4	1	AAP10133	Sequence
287	5	55.6	3	6	ABU19424	Neural th	360	5	55.6	4	1	AAP10129	Sequence
288	5	55.6	3	6	ABU06547	Maize sta	361	5	55.6	4	1	AAP10138	Sequence
289	5	55.6	3	6	ABU12339	Angiotens	362	5	55.6	4	1	AAP10130	Sequence
290	5	55.6	3	6	ABR43883	Synovial	363	5	55.6	4	1	AAP10135	Sequence
291	5	55.6	3	6	ABR43890	Synovial	364	5	55.6	4	1	AAP10132	Sequence
292	5	55.6	3	6	ABR43889	Synovial	365	5	55.6	4	1	AAP10134	Sequence
293	5	55.6	3	6	ABR43891	Synovial	366	5	55.6	4	1	AAP10136	Sequence
294	5	55.6	3	6	ABO44788	Novel hum	367	5	55.6	4	1	AAP10131	Sequence
295	5	55.6	3	6	ABR82577	Protein k	368	5	55.6	4	1	AAP10139	Sequence
296	5	55.6	3	6	ABR82578	Protein k	369	5	55.6	4	1	AAP10365	Peptide s
297	5	55.6	3	6	ABR82579	Protein k	370	5	55.6	4	1	AAP10364	Peptide s
298	5	55.6	3	6	ABR82606	Protein k	371	5	55.6	4	1	AAP10362	Peptide s
299	5	55.6	3	6	ABR82580	Protein k	372	5	55.6	4	1	AAP10078	Sequence
300	5	55.6	3	7	ABO26268	Fragment	373	5	55.6	4	1	AAP10080	Sequence
301	5	55.6	3	7	ABO26268	Human pro	374	5	55.6	4	1	AAP10092	Sequence
302	5	55.6	3	7	ADC36727	Mutated v	375	5	55.6	4	1	AAP10079	Sequence
303	5	55.6	3	7	ADC18408	Protease	376	5	55.6	4	1	AAP10093	Sequence
304	5	55.6	3	7	ADC37014	Cell bind	377	5	55.6	4	1	AAP10094	Sequence
305	5	55.6	3	7	ADP10724	Human sec	378	5	55.6	4	1	AAP20247	Hydrolysa
306	5	55.6	3	7	ADK41651	Bradykini	379	5	55.6	4	1	AAP20258	Protected
307	5	55.6	3	7	ADW33443	HLA bindi	380	5	55.6	4	1	AAP20062	Tuftsin
308	5	55.6	3	7	ADW37077	HLA bindi	381	5	55.6	4	1	AAP20059	Enzyme su
309	5	55.6	3	7	ADW36789	HLA bindi	382	5	55.6	4	1	AAP30082	Sequence
310	5	55.6	3	7	ADW35709	HLA bindi	383	5	55.6	4	1	AAP30083	Sequence
311	5	55.6	3	7	ADW34591	HLA bindi	384	5	55.6	4	1	AAP30085	Sequence
312	5	55.6	3	7	ADW36683	HLA bindi	385	5	55.6	4	1	AAP30084	Sequence
313	5	55.6	3	7	ADW33358	HLA bindi	386	5	55.6	4	1	AAP30132	Sequence
314	5	55.6	3	7	ADW34676	HLA bindi	387	5	55.6	4	1	AAP30127	Sequence
315	5	55.6	3	7	ADW36739	HLA bindi	388	5	55.6	4	1	AAP30130	Sequence
316	5	55.6	3	7	ADW34476	HLA bindi	389	5	55.6	4	1	AAP30129	Sequence

390	5	55.6	4	1	AAP30131	Sequence	463	5	55.6	4	1	AAP91418	Sequence
391	5	55.6	4	1	AAP30133	Sequence	464	5	55.6	4	1	AAP91620	PNS/CNS m
392	5	55.6	4	1	AAP30128	Sequence	465	5	55.6	4	1	AAP91630	PNS/CNS m
393	5	55.6	4	1	AAP30126	Sequence	466	5	55.6	4	1	AAP91610	Motif use
394	5	55.6	4	1	AAP30135	Sequence	467	5	55.6	4	1	AAP91624	Motif use
395	5	55.6	4	1	AAP30088	Mammalian	468	5	55.6	4	1	AAP91627	Motif use
396	5	55.6	4	1	AAP30087	Mammalian	469	5	55.6	4	1	AAP93304	Anti-thro
397	5	55.6	4	1	AAP40482	Substrate	470	5	55.6	4	1	AAP91921	Anti-thro
398	5	55.6	4	1	AAP40562	Oligopept	471	5	55.6	4	1	AAP92312	Fibronect
399	5	55.6	4	1	AAP40562	Oligopept	472	5	55.6	4	1	AAP92312	Fibronect
400	5	55.6	4	1	AAP40456	N-termina	473	5	55.6	4	2	AAR05705	Peptide h
401	5	55.6	4	1	AAP40461	N-termina	474	5	55.6	4	2	AAR05704	Peptide h
402	5	55.6	4	1	AAP40464	N-termina	475	5	55.6	4	2	AAR05702	Peptide h
403	5	55.6	4	1	AAP40467	Enkephali	476	5	55.6	4	2	AAR03676	Peptide h
404	5	55.6	4	1	AAP40354	Sequence	477	5	55.6	4	2	AAR03677	Peptide h
405	5	55.6	4	1	AAP50690	Analgesic	478	5	55.6	4	2	AAR05706	Peptide h
406	5	55.6	4	1	AAP50336	Analgesic	479	5	55.6	4	2	AAR03675	Peptide h
407	5	55.6	4	1	AAP50109	N-termina	480	5	55.6	4	2	AAR07429	Vasopress
408	5	55.6	4	1	AAP50542	Synthetic	481	5	55.6	4	2	AAR07428	Vasopress
409	5	55.6	4	1	AAP50015	Sequence	482	5	55.6	4	2	AAR06633	Peptide c
410	5	55.6	4	1	AAP50016	Sequence	483	5	55.6	4	2	AAR06633	Peptide c
411	5	55.6	4	1	AAP61324	Sequence	484	5	55.6	4	2	AAR04608	Antiviral
412	5	55.6	4	1	AAP60484	Peptide w	485	5	55.6	4	2	AAR05320	Anti-coag
413	5	55.6	4	1	AAP60497	Peptide w	486	5	55.6	4	2	AAR05860	Anti-coag
414	5	55.6	4	1	AAP60848	Hypotensi	487	5	55.6	4	2	AAR07884	Peptide c
415	5	55.6	4	1	AAP60850	Hypotensi	488	5	55.6	4	2	AAR07103	Melanocyt
416	5	55.6	4	1	AAP60851	Hypotensi	489	5	55.6	4	2	AAR07102	Melanocyt
417	5	55.6	4	1	AAP60854	Hypotensi	490	5	55.6	4	2	AAR05919	Low toxic
418	5	55.6	4	1	AAP60855	Hypotensi	491	5	55.6	4	2	AAR05916	Low toxic
419	5	55.6	4	1	AAP60852	Hypotensi	492	5	55.6	4	2	AAR05917	Low toxic
420	5	55.6	4	1	AAP61666	Sequence	493	5	55.6	4	2	AAR15424	Coagulati
421	5	55.6	4	1	AAP61667	Sequence	494	5	55.6	4	2	AAR12817	Reagin-inh
422	5	55.6	4	1	AAP61669	Sequence	495	5	55.6	4	2	AAR12817	Reagin-inh
423	5	55.6	4	1	AAP61358	Sequence	496	5	55.6	4	2	AAR14267	Chromogen
424	5	55.6	4	1	AAP61356	Sequence	497	5	55.6	4	2	AAR11744	Cyclic pl
425	5	55.6	4	1	AAP61354	Sequence	498	5	55.6	4	2	AAR11745	Cyclic pl
426	5	55.6	4	1	AAP61491	Analgesic	499	5	55.6	4	2	AAR13257	Cytotoxic
427	5	55.6	4	1	AAP61741	Analgesic	500	5	55.6	4	2	AAR13258	Cytotoxic
428	5	55.6	4	1	AAP60157	Human gro	501	5	55.6	4	2	AAR12870	Chromogen
429	5	55.6	4	1	AAP61325	Sequence	502	5	55.6	4	2	AAR15327	Tuftsina a
430	5	55.6	4	1	AAP711341	Analgesic	503	5	55.6	4	2	AAR14433	Tuftsina a
431	5	55.6	4	1	AAP711700	Analgesic	504	5	55.6	4	2	AAR14947	Tuftsina a
432	5	55.6	4	1	AAP711484	Sequence	505	5	55.6	4	2	AAR10408	Fibrinoge
433	5	55.6	4	1	AAP711482	Sequence	506	5	55.6	4	2	AAR11750	Platelet
434	5	55.6	4	1	AAP711483	Sequence	507	5	55.6	4	2	AAR11752	Platelet
435	5	55.6	4	1	AAP71580	Growth ho	508	5	55.6	4	2	AAR11751	Platelet
436	5	55.6	4	1	AAP71582	Growth ho	509	5	55.6	4	2	AAR12407	Anti-thro
437	5	55.6	4	1	AAP71579	Growth ho	510	5	55.6	4	2	AAR13808	Factor Xa
438	5	55.6	4	1	AAP71602	Hypotensi	511	5	55.6	4	2	AAR13807	Factor Xa
439	5	55.6	4	1	AAP71604	Hypotensi	512	5	55.6	4	2	AAR13810	Factor Xa
440	5	55.6	4	1	AAP71607	Hypotensi	513	5	55.6	4	2	AAR13803	Factor Xa
441	5	55.6	4	1	AAP71601	Hypotensi	514	5	55.6	4	2	AAR13805	Factor Xa
442	5	55.6	4	1	AAP71605	Hypotensi	515	5	55.6	4	2	AAR13804	Factor Xa
443	5	55.6	4	1	AAP71608	Hypotensi	516	5	55.6	4	2	AAR13812	Factor Xa
444	5	55.6	4	1	AAP80873	Hook regi	517	5	55.6	4	2	AAR13806	Factor Xa
445	5	55.6	4	1	AAP82910	Acetylcho	518	5	55.6	4	2	AAR14686	Hyaluroni
446	5	55.6	4	1	AAP80424	Sequence	519	5	55.6	4	2	AAR27101	Sequence
447	5	55.6	4	1	AAP80994	Sequence	520	5	55.6	4	2	AAR22519	Peptide o
448	5	55.6	4	1	AAP83217	Sequence	521	5	55.6	4	2	AAR22525	Peptide o
449	5	55.6	4	1	AAP82678	Thymopoie	522	5	55.6	4	2	AAR22518	Peptide o
450	5	55.6	4	1	AAB62306	Factor Xa	523	5	55.6	4	2	AAR22527	Peptide o
451	5	55.6	4	1	AAP80258	Sequence	524	5	55.6	4	2	AAR22520	Peptide o
452	5	55.6	4	1	AAP93521	Amino aci	525	5	55.6	4	2	AAR22521	Peptide o
453	5	55.6	4	1	AAP93594	Peptide a	526	5	55.6	4	2	AAR22522	Peptide o
454	5	55.6	4	1	AAP91371	Peptide a	527	5	55.6	4	2	AAR22524	Peptide o
455	5	55.6	4	1	AAP93313	Peptide a	528	5	55.6	4	2	AAR27701	Peptide o
456	5	55.6	4	1	AAP93596	Peptide a	529	5	55.6	4	2	AAR22517	Peptide o
457	5	55.6	4	1	AAP93599	Peptide a	530	5	55.6	4	2	AAR22526	Peptide o
458	5	55.6	4	1	AAP93597	Peptide a	531	5	55.6	4	2	AAR22410	CPase B.1
459	5	55.6	4	1	AAP93312	Peptide a	532	5	55.6	4	2	AAR27340	Allopurin
460	5	55.6	4	1	AAP93311	Peptide a	533	5	55.6	4	2	AAR27339	Allopurin
461	5	55.6	4	1	AAP93595	Peptide a	534	5	55.6	4	2	AAR22267	Protectiv
462	5	55.6	4	1	AAP93598	Peptide a	535	5	55.6	4	2	AAR28903	RGase par



536	5	55.6	4	2	AAR233935	Aar23935	Lactoferr	609	5	55.6	4	2	AAR32386	Aar32386	Fibrinoge
537	5	55.6	4	2	AAR29783	Aar29783	Selective	610	5	55.6	4	2	AAR32385	Aar32385	Fibrinoge
538	5	55.6	4	2	AAR22992	Aar22992	Glucosamyl	611	5	55.6	4	2	AAR32377	Aar32377	Fibrinoge
539	5	55.6	4	2	AAR26802	Aar26802	Triglycer	612	5	55.6	4	2	AAR36616	Aar36616	Group I s
540	5	55.6	4	2	AAR24214	Aar24214	Fragment	613	5	55.6	4	2	AAR37833	Aar37833	Cell adhe
541	5	55.6	4	2	AAR24213	Aar24213	Fragment	614	5	55.6	4	2	AAR38743	Aar38743	KEX2 clea
542	5	55.6	4	2	AAR31243	Aar31243	HIV princ	615	5	55.6	4	2	AAR38742	Aar38742	KEX2 clea
543	5	55.6	4	2	AAR27138	Aar27138	Fibronect	616	5	55.6	4	2	AAR32109	Aar32109	Sequence
544	5	55.6	4	2	AAR27136	Aar27136	Fibronect	617	5	55.6	4	2	AAR26267	Aar26267	Cysteine
545	5	55.6	4	2	AAR27135	Aar27135	Fibronect	618	5	55.6	4	2	AAR26268	Aar26268	Cysteine
546	5	55.6	4	2	AAR27137	Aar27137	Fibronect	619	5	55.6	4	2	AAR43092	Aar43092	Diuretic
547	5	55.6	4	2	AAR27134	Aar27134	Fibronect	620	5	55.6	4	2	AAR43091	Aar43091	Diuretic
548	5	55.6	4	2	AAR32270	Aar32270	Soybean g	621	5	55.6	4	2	AAR43089	Aar43089	Diuretic
549	5	55.6	4	2	AAR26423	Aar26423	Immune en	622	5	55.6	4	2	AAR43100	Aar43100	Diuretic
550	5	55.6	4	2	AAR26422	Aar26422	Immune en	623	5	55.6	4	2	AAR44043	Aar44043	RGD pepri
551	5	55.6	4	2	AAR26421	Aar26421	Immune en	624	5	55.6	4	2	AAR41638	Aar41638	Internal
552	5	55.6	4	2	AAR26424	Aar26424	Immune en	625	5	55.6	4	2	AAR41649	Aar41649	Internal
553	5	55.6	4	2	AAR23726	Aar23726	ACE inhib	626	5	55.6	4	2	AAR41650	Aar41650	Internal
554	5	55.6	4	2	AAR22935	Aar22935	Neurotrop	627	5	55.6	4	2	AAR41643	Aar41643	Internal
555	5	55.6	4	2	AAR22936	Aar22936	Neurotrop	628	5	55.6	4	2	AAR41642	Aar41642	Internal
556	5	55.6	4	2	AAR22937	Aar22937	Neurotrop	629	5	55.6	4	2	AAR41636	Aar41636	Internal
557	5	55.6	4	2	AAR25311	Aar25311	Cell cont	630	5	55.6	4	2	AAR41637	Aar41637	Internal
558	5	55.6	4	2	AAR25314	Aar25314	Cell cont	631	5	55.6	4	2	AAR32651	Aar32651	Substance
559	5	55.6	4	2	AAR25315	Aar25315	Cell cont	632	5	55.6	4	2	AAR32652	Aar32652	C-termina
560	5	55.6	4	2	AAR25318	Aar25318	Cell cont	633	5	55.6	4	2	AAR32653	Aar32653	N-termina
561	5	55.6	4	2	AAR25313	Aar25313	Cell cont	634	5	55.6	4	2	AAR39390	Aar39390	Factor Xa
562	5	55.6	4	2	AAR25312	Aar25312	Cell cont	635	5	55.6	4	2	AAR37134	Aar37134	RGD pepri
563	5	55.6	4	2	AAR25316	Aar25316	Cell cont	636	5	55.6	4	2	AAR37133	Aar37133	RGD pepri
564	5	55.6	4	2	AAR25317	Aar25317	Cell cont	637	5	55.6	4	2	AAR37137	Aar37137	RGD pepri
565	5	55.6	4	2	AAR30046	Aar30046	Cyclic HI	638	5	55.6	4	2	AAR37138	Aar37138	RGD pepri
566	5	55.6	4	2	AAR31311	Aar31311	Alpha-sub	639	5	55.6	4	2	AAR37132	Aar37132	RGD pepri
567	5	55.6	4	2	AAR31312	Aar31312	Alpha-sub	640	5	55.6	4	2	AAR37135	Aar37135	RGD pepri
568	5	55.6	4	2	AAR33016	Aar33016	Alpha-sub	641	5	55.6	4	2	AAR37130	Aar37130	RGD pepri
569	5	55.6	4	2	AAR21665	Aar21665	Cyclic te	642	5	55.6	4	2	AAR37136	Aar37136	RGD pepri
570	5	55.6	4	2	AAR21664	Aar21664	Cyclic te	643	5	55.6	4	2	AAR43515	Aar43515	Sm B/B' e
571	5	55.6	4	2	AAR23907	Aar23907	Peptide 1	644	5	55.6	4	2	AAR31392	Aar31392	Cyclic pl
572	5	55.6	4	2	AAR25241	Aar25241	PTHp (10	645	5	55.6	4	2	AAR41987	Aar41987	Polymorph
573	5	55.6	4	2	AAR20131	Aar20131	SEQ ID No	646	5	55.6	4	2	AAR34873	Aar34873	Endotheli
574	5	55.6	4	2	AAR26810	Aar26810	Polyethyl	647	5	55.6	4	2	AAR34874	Aar34874	Endotheli
575	5	55.6	4	2	AAR26812	Aar26812	Polyethyl	648	5	55.6	4	2	AAR69323	Aar69323	Gp IIB/II
576	5	55.6	4	2	AAR25457	Aar25457	woh1-4_3	649	5	55.6	4	2	AAR69324	Aar69324	Gp IIB/II
577	5	55.6	4	2	AAR27863	Aar27863	Antimicro	650	5	55.6	4	2	AAR38485	Aar38485	Human RDS
578	5	55.6	4	2	AAR23084	Aar23084	Fluorinat	651	5	55.6	4	2	AAR33343	Aar33343	Sequence
579	5	55.6	4	2	AAR26806	Aar26806	Propen-am	652	5	55.6	4	2	AAR33292	Aar33292	FGF antag
580	5	55.6	4	2	AAR26394	Aar26394	Sequence	653	5	55.6	4	2	AAR35464	Aar35464	Propene-a
581	5	55.6	4	2	AAR26392	Aar26392	Sequence	654	5	55.6	4	2	AAR35467	Aar35467	Propene-a
582	5	55.6	4	2	AAR26398	Aar26398	Sequence	655	5	55.6	4	2	AAR37588	Aar37588	Factor Xa
583	5	55.6	4	2	AAR21345	Aar21345	HIV prote	656	5	55.6	4	2	AAR32185	Aar32185	Proteolyc
584	5	55.6	4	2	AAR21000	Aar21000	HIV prote	657	5	55.6	4	2	AAR38341	Aar38341	Diuretic
585	5	55.6	4	2	AAR20174	Aar20174	Hypoxanth	658	5	55.6	4	2	AAR38342	Aar38342	Cell adhe
586	5	55.6	4	2	AAR25999	Aar25999	Tetrapt	659	5	55.6	4	2	AAY18003	Aay18003	Cell adhe
587	5	55.6	4	2	AAR26678	Aar26678	HIV-pND-p	660	5	55.6	4	2	AAR32189	Aar32189	Serine pr
588	5	55.6	4	2	AAR36713	Aar36713	Adhesion	661	5	55.6	4	2	AAR42570	Aar42570	Peptide C
589	5	55.6	4	2	AAR43239	Aar43239	Deletion	662	5	55.6	4	2	AAR34514	Aar34514	Ala-Ala-G
590	5	55.6	4	2	AAR33662	Aar33662	Growth ho	663	5	55.6	4	2	AAR34512	Aar34512	Arg-Tyr-H
591	5	55.6	4	2	AAR33660	Aar33660	Growth ho	664	5	55.6	4	2	AAR38101	Aar38101	Protease-
592	5	55.6	4	2	AAR33664	Aar33664	Growth ho	665	5	55.6	4	2	AAR38116	Aar38116	Protease-
593	5	55.6	4	2	AAR39732	Aar39732	First typ	666	5	55.6	4	2	AAR38132	Aar38132	Protease-
594	5	55.6	4	2	AAR37634	Aar37634	Sequence	667	5	55.6	4	2	AAR38106	Aar38106	Protease-
595	5	55.6	4	2	AAR37635	Aar37635	Sequence	668	5	55.6	4	2	AAR38108	Aar38108	Protease-
596	5	55.6	4	2	AAR37630	Aar37630	Sequence	669	5	55.6	4	2	AAR38124	Aar38124	Protease-
597	5	55.6	4	2	AAR37628	Aar37628	Sequence	670	5	55.6	4	2	AAR38137	Aar38137	Protease-
598	5	55.6	4	2	AAR37633	Aar37633	Sequence	671	5	55.6	4	2	AAR38128	Aar38128	Protease-
599	5	55.6	4	2	AAR37627	Aar37627	Sequence	672	5	55.6	4	2	AAR38109	Aar38109	Protease-
600	5	55.6	4	2	AAR37632	Aar37632	Sequence	673	5	55.6	4	2	AAR38136	Aar38136	Protease-
601	5	55.6	4	2	AAR37629	Aar37629	Sequence	674	5	55.6	4	2	AAR38126	Aar38126	Protease-
602	5	55.6	4	2	AAR37631	Aar37631	Sequence	675	5	55.6	4	2	AAR38145	Aar38145	Protease-
603	5	55.6	4	2	AAR36523	Aar36523	C-amidate	676	5	55.6	4	2	AAR38385	Aar38385	Cyclic HI
604	5	55.6	4	2	AAR32383	Aar32383	Fibrinoge	677	5	55.6	4	2	AAR43069	Aar43069	HVD large
605	5	55.6	4	2	AAR32382	Aar32382	Fibrinoge	678	5	55.6	4	2	AAR52978	Aar52978	Tetrapt
606	5	55.6	4	2	AAR32381	Aar32381	Fibrinoge	679	5	55.6	4	2	AAR48072	Aar48072	Biologica
607	5	55.6	4	2	AAR32380	Aar32380	Fibrinoge	680	5	55.6	4	2	AAR48071	Aar48071	Biologica
608	5	55.6	4	2	AAR32384	Aar32384	Fibrinoge	681	5	55.6	4	2	AAR48079	Aar48079	Intra/Int

682	5	55.6	4	2	AAR60693	H. haliobi	755	5	55.6	4	2	AAR57449	Lactoferr
683	5	55.6	4	2	AAR48252	Endoplasm	756	5	55.6	4	2	AAR57453	Lactoferr
684	5	55.6	4	2	AAR48281	Nuclear r	757	5	55.6	4	2	AAR57457	Lactoferr
685	5	55.6	4	2	AAR57422	Rabphilin	758	5	55.6	4	2	AAR5804	Oligopept
686	5	55.6	4	2	AAR46480	Fragment	759	5	55.6	4	2	AAR65802	Oligopept
687	5	55.6	4	2	AAR46803	Phytase d	760	5	55.6	4	2	AAR56760	C-terminu
688	5	55.6	4	2	AAR63104	Inflammat	761	5	55.6	4	2	AAR57899	Integrin
689	5	55.6	4	2	AAR58805	Interleuk	762	5	55.6	4	2	AAR57900	Integrin
690	5	55.6	4	2	AAR4362	Sequence	763	5	55.6	4	2	AAR57885	Human vit
691	5	55.6	4	2	AAR53589	Pig somat	764	5	55.6	4	2	AAR57898	Integrin
692	5	55.6	4	2	AAR42561	Tri-/tetr	765	5	55.6	4	2	AAR57897	Integrin
693	5	55.6	4	2	AAR42558	Tri-/tetr	766	5	55.6	4	2	AAR48599	Sequence
694	5	55.6	4	2	AAR42568	Tri-/tetr	767	5	55.6	4	2	AAR48599	Sequence
695	5	55.6	4	2	AAR55421	Opioid ol	768	5	55.6	4	2	AAR61350	Vasopress
696	5	55.6	4	2	AAR55441	Opioid ol	769	5	55.6	4	2	AAR49797	Sequence
697	5	55.6	4	2	AAR55469	Opioid ol	770	5	55.6	4	2	AAR49795	Sequence
698	5	55.6	4	2	AAR55502	Opioid ol	771	5	55.6	4	2	AAR49800	Sequence
699	5	55.6	4	2	AAR55435	Opioid ol	772	5	55.6	4	2	AAR49798	Sequence
700	5	55.6	4	2	AAR55481	Opioid ol	773	5	55.6	4	2	AAR49796	Sequence
701	5	55.6	4	2	AAR55497	Opioid ol	774	5	55.6	4	2	AAR49799	Sequence
702	5	55.6	4	2	AAR55503	Opioid ol	775	5	55.6	4	2	AAR49794	Sequence
703	5	55.6	4	2	AAR55455	Opioid ol	776	5	55.6	4	2	AAR49793	Sequence
704	5	55.6	4	2	AAR55465	Opioid ol	777	5	55.6	4	2	AAR60559	Tissue pl
705	5	55.6	4	2	AAR55494	Opioid ol	778	5	55.6	4	2	AAR64131	Memory-st
706	5	55.6	4	2	AAR55507	Opioid ol	779	5	55.6	4	2	AAR59927	Alpha 1-a
707	5	55.6	4	2	AAR55417	Opioid ol	780	5	55.6	4	2	AAR54863	Factor Xa
708	5	55.6	4	2	AAR55431	Opioid ol	781	5	55.6	4	2	AAR45689	Selectin
709	5	55.6	4	2	AAR55425	Opioid ol	782	5	55.6	4	2	AAR45689	Peptide d
710	5	55.6	4	2	AAR55498	Opioid ol	783	5	55.6	4	2	AAR66127	Peptide d
711	5	55.6	4	2	AAR55419	Opioid ol	784	5	55.6	4	2	AAR48522	Lactoferr
712	5	55.6	4	2	AAR55433	Opioid ol	785	5	55.6	4	2	AAR48526	Lactoferr
713	5	55.6	4	2	AAR55451	Opioid ol	786	5	55.6	4	2	AAR48518	Lactoferr
714	5	55.6	4	2	AAR55477	Opioid ol	787	5	55.6	4	2	AAR48521	Lactoferr
715	5	55.6	4	2	AAR55506	Opioid ol	788	5	55.6	4	2	AAR63967	C5a anaph
716	5	55.6	4	2	AAR55445	Opioid ol	789	5	55.6	4	2	AAR63976	C5a anaph
717	5	55.6	4	2	AAR55475	Opioid ol	790	5	55.6	4	2	AAR63971	C5a anaph
718	5	55.6	4	2	AAR55449	Opioid ol	791	5	55.6	4	2	AAR63963	C5a anaph
719	5	55.6	4	2	AAR55493	Opioid ol	792	5	55.6	4	2	AAR64001	C5a anaph
720	5	55.6	4	2	AAR55447	Opioid ol	793	5	55.6	4	2	AAR63973	C5a anaph
721	5	55.6	4	2	AAR55484	Opioid ol	794	5	55.6	4	2	AAR63969	C5a anaph
722	5	55.6	4	2	AAR55415	Opioid ol	795	5	55.6	4	2	AAR63979	C5a anaph
723	5	55.6	4	2	AAR55437	Opioid ol	796	5	55.6	4	2	AAR51936	Helix pom
724	5	55.6	4	2	AAR55443	Opioid ol	797	5	55.6	4	2	AAR51932	Helix pom
725	5	55.6	4	2	AAR55423	Opioid ol	798	5	55.6	4	2	AAR60592	Native pr
726	5	55.6	4	2	AAR55439	Opioid ol	799	5	55.6	4	2	AAR60595	proins.KQ
727	5	55.6	4	2	AAR55485	Opioid ol	800	5	55.6	4	2	AAR60591	Native pr
728	5	55.6	4	2	AAR55488	Opioid ol	801	5	55.6	4	2	AAR60593	proins.RT
729	5	55.6	4	2	AAR55453	Opioid ol	802	5	55.6	4	2	AAR61198	Macrophag
730	5	55.6	4	2	AAR55471	Opioid ol	803	5	55.6	4	2	AAR51222	Platelet
731	5	55.6	4	2	AAR55479	Opioid ol	804	5	55.6	4	2	AAR51221	Platelet
732	5	55.6	4	2	AAR55467	Opioid ol	805	5	55.6	4	2	AAR51493	Human A-m
733	5	55.6	4	2	AAR55473	Opioid ol	806	5	55.6	4	2	AAR51482	SV40 T an
734	5	55.6	4	2	AAR55489	Opioid ol	807	5	55.6	4	2	AAR51484	Dorsal NL
735	5	55.6	4	2	AAR60534	Peptide c	808	5	55.6	4	2	AAR51485	Human c-m
736	5	55.6	4	2	AAR60533	Peptide c	809	5	55.6	4	2	AAR51485	Human c-m
737	5	55.6	4	2	AAR51007	Human HSP	810	5	55.6	4	2	AAR51496	Dorsal cd
738	5	55.6	4	2	AAR44665	Platelet	811	5	55.6	4	2	AAR51491	Human A-m
739	5	55.6	4	2	AAR44663	Platelet	812	5	55.6	4	2	AAR54528	Platelet
740	5	55.6	4	2	AAR44667	Platelet	813	5	55.6	4	2	AAR54526	Platelet
741	5	55.6	4	2	AAR44662	Platelet	814	5	55.6	4	2	AAR54527	Platelet
742	5	55.6	4	2	AAR67015	RGD pepti	815	5	55.6	4	2	AAR54529	Platelet
743	5	55.6	4	2	AAR56656	Platelet	816	5	55.6	4	2	AAR60327	Blood coa
744	5	55.6	4	2	AAR56655	Platelet	817	5	55.6	4	2	AAR54662	Native se
745	5	55.6	4	2	AAR56654	Platelet	818	5	55.6	4	2	AAR54662	Native se
746	5	55.6	4	2	AAR47431	ACE inhib	819	5	55.6	4	2	AAR54529	Tetrapept
747	5	55.6	4	2	AAR57253	Cyclized	820	5	55.6	4	2	AAR55549	Tetrapept
748	5	55.6	4	2	AAR62155	Basic/aci	821	5	55.6	4	2	AAR67594	Tetrapept
749	5	55.6	4	2	AAR51593	Peptide a	822	5	55.6	4	2	AAR67594	Tetrapept
750	5	55.6	4	2	AAR51589	Peptide a	823	5	55.6	4	2	AAR67594	Tetrapept
751	5	55.6	4	2	AAR51591	Peptide a	824	5	55.6	4	2	AAR48940	Thrombin
752	5	55.6	4	2	AAR51595	Peptide a	825	5	55.6	4	2	AAR48942	Thrombin
753	5	55.6	4	2	AAR47186	Angiotens	826	5	55.6	4	2	AAR48942	Thrombin
754	5	55.6	4	2	AAR57452	Lactoferr	827	5	55.6	4	2	AAR54581	Cholecyst

828	5	55.6	4	2	AAR54585	Cholécyst	901	5	55.6	4	2	AAR82448	Optional
829	5	55.6	4	2	AAR54580	Cholécyst	902	5	55.6	4	2	AAR76136	hML(152-1
830	5	55.6	4	2	AAR54573	Cholécyst	903	5	55.6	4	2	AAR88292	Memory en
831	5	55.6	4	2	AAR54570	Cholécyst	904	5	55.6	4	2	AAR61396	PF4-relat
832	5	55.6	4	2	AAR54582	Cholécyst	905	5	55.6	4	2	AAR88219	Lactoferr
833	5	55.6	4	2	AAR54571	Cholécyst	906	5	55.6	4	2	AAR86902	Thrombin
834	5	55.6	4	2	AAR54586	Cholécyst	907	5	55.6	4	2	AAR86901	Blood coa
835	5	55.6	4	2	AAR54572	Cholécyst	908	5	55.6	4	2	AAR87659	HIS- (D) Ph
836	5	55.6	4	2	AAR69469	Integrin-	909	5	55.6	4	2	AAR87663	HIS- (D) Ph
837	5	55.6	4	2	AAR69474	Integrin-	910	5	55.6	4	2	AAR77843	RGDV pept
838	5	55.6	4	2	AAR42585	Kyotórphi	911	5	55.6	4	2	AAR66944	RACE pept
839	5	55.6	4	2	AAR42586	Kyotórphi	912	5	55.6	4	2	AAR64690	Bovine la
840	5	55.6	4	2	AAR42587	Kyotórphi	913	5	55.6	4	2	AAR84686	Bovine la
841	5	55.6	4	2	AAR42588	Kyotórphi	914	5	55.6	4	2	AAR84689	Bovine la
842	5	55.6	4	2	AAR59394	Thy spleni	915	5	55.6	4	2	AAR84694	Bovine la
843	5	55.6	4	2	AAR59401	Thy spleni	916	5	55.6	4	2	AAR83116	Factor-Xa
844	5	55.6	4	2	AAR59396	Thy spleni	917	5	55.6	4	2	AAR09862	Thrombin
845	5	55.6	4	2	AAR59390	Thy spleni	918	5	55.6	4	2	AAR09861	Thrombin
846	5	55.6	4	2	AAR59402	Thy spleni	919	5	55.6	4	2	AAR68591	Rat NDF p
847	5	55.6	4	2	AAR59389	Thy spleni	920	5	55.6	4	2	AAR80049	Peptidase
848	5	55.6	4	2	AAR59404	Thy spleni	921	5	55.6	4	2	AAR80056	Peptidase
849	5	55.6	4	2	AAR59393	Thy spleni	922	5	55.6	4	2	AAR83236	Integrin-
850	5	55.6	4	2	AAR59391	Thy spleni	923	5	55.6	4	2	AAR83231	Integrin-
851	5	55.6	4	2	AAR59398	Thy spleni	924	5	55.6	4	2	AAR87691	Integrin-
852	5	55.6	4	2	AAR59397	Thy spleni	925	5	55.6	4	2	AAR87690	Oligopept
853	5	55.6	4	2	AAR59409	Thy spleni	926	5	55.6	4	2	AAR87695	Oligopept
854	5	55.6	4	2	AAR59395	Thy spleni	927	5	55.6	4	2	AAR87694	Oligopept
855	5	55.6	4	2	AAR59403	Thy spleni	928	5	55.6	4	2	AAR87689	Oligopept
856	5	55.6	4	2	AAR59399	Thy spleni	929	5	55.6	4	2	AAR79757	Anti-para
857	5	55.6	4	2	AAR59408	Thy spleni	930	5	55.6	4	2	AAR80260	Anti-para
858	5	55.6	4	2	AAR59392	Thy spleni	931	5	55.6	4	2	AAR79759	Anti-para
859	5	55.6	4	2	AAR59400	Thy spleni	932	5	55.6	4	2	AAR79760	Anti-para
860	5	55.6	4	2	AAR57828	RGD contg	933	5	55.6	4	2	AAR75601	gp120 bin
861	5	55.6	4	2	AAR60506	Factor Xa	934	5	55.6	4	2	AAR75603	gp120 bin
862	5	55.6	4	2	AAR60511	Factor Xa	935	5	55.6	4	2	AAR69104	Activated
863	5	55.6	4	2	AAR60510	Factor Xa	936	5	55.6	4	2	AAR85667	Anti-alle
864	5	55.6	4	2	AAR60507	Factor Xa	937	5	55.6	4	2	AAR80463	Artificia
865	5	55.6	4	2	AAR60509	Factor Xa	938	5	55.6	4	2	AAR80464	Artificia
866	5	55.6	4	2	AAR60504	Factor Xa	939	5	55.6	4	2	AAR79354	Human con
867	5	55.6	4	2	AAR60508	Factor Xa	940	5	55.6	4	2	AAR79357	Mutant th
868	5	55.6	4	2	AAR60512	Factor Xa	941	5	55.6	4	2	AAR79363	Mutant th
869	5	55.6	4	2	AAR52965	Human A-m	942	5	55.6	4	2	AAR79361	Mutant th
870	5	55.6	4	2	AAR52956	SV40: T an	943	5	55.6	4	2	AAR79360	Mutant th
871	5	55.6	4	2	AAR52968	Dorsal NL	944	5	55.6	4	2	AAR69769	Thrombos
872	5	55.6	4	2	AAR52959	Human C-m	945	5	55.6	4	2	AAR79362	Mutant th
873	5	55.6	4	2	AAR52970	Dorsal cd	946	5	55.6	4	2	AAR69768	Thrombos
874	5	55.6	4	2	AAR52967	Human A-m	947	5	55.6	4	2	AAR69767	Thrombos
875	5	55.6	4	2	AAR55046	161 legio	948	5	55.6	4	2	AAR79356	Thrombos
876	5	55.6	4	2	AAR55042	161 legio	949	5	55.6	4	2	AAR69766	Thrombos
877	5	55.6	4	2	AAR53715	Tetrápept	950	5	55.6	4	2	AAR62945	RGD contg
878	5	55.6	4	2	AAR53728	Tetrápept	951	5	55.6	4	2	AAR62947	RGD contg
879	5	55.6	4	2	AAR53719	Tetrápept	952	5	55.6	4	2	AAR62946	RGD contg
880	5	55.6	4	2	AAR53712	Tetrápept	953	5	55.6	4	2	AAR08853	Peptide c
881	5	55.6	4	2	AAR53722	Tetrápept	954	5	55.6	4	2	AAR08855	Peptide c
882	5	55.6	4	2	AAR53717	Tetrápept	955	5	55.6	4	2	AAR08854	Peptide c
883	5	55.6	4	2	AAR53721	Tetrápept	956	5	55.6	4	2	AAR64811	ScFv-Lys
884	5	55.6	4	2	AAR53724	Tetrápept	957	5	55.6	4	2	AAR65180	Dibasic a
885	5	55.6	4	2	AAR53718	Tetrápept	958	5	55.6	4	2	AAR89791	Melanotro
886	5	55.6	4	2	AAR53723	Tetrápept	959	5	55.6	4	2	AAR89788	Melanotro
887	5	55.6	4	2	AAR53726	Tetrápept	960	5	55.6	4	2	AAR81988	Peptidyl
888	5	55.6	4	2	AAR53729	Tetrápept	961	5	55.6	4	2	AAR84895	ATPR2Ap
889	5	55.6	4	2	AAR65261	Peptide c	962	5	55.6	4	2	AAR67881	HIV virus
890	5	55.6	4	2	AAR89742	C5a C-ter	963	5	55.6	4	2	AAR67887	HIV virus
891	5	55.6	4	2	AAR89743	C5a C-ter	964	5	55.6	4	2	AAR67880	HIV virus
892	5	55.6	4	2	AAR89741	C5a C-ter	965	5	55.6	4	2	AAR89771	Synthetic
893	5	55.6	4	2	AAR89740	C5a C-ter	966	5	55.6	4	2	AAR89772	Synthetic
894	5	55.6	4	2	AAR89739	C5a C-ter	967	5	55.6	4	2	AAR82908	Non-RGD
895	5	55.6	4	2	AAR13702	Factor Xa	968	5	55.6	4	2	AAR70473	Cancer me
896	5	55.6	4	2	AAR85941	Peptidase	969	5	55.6	4	2	AAR70476	Cancer me
897	5	55.6	4	2	AAR85940	Peptidase	970	5	55.6	4	2	AAR72663	Cladospor
898	5	55.6	4	2	AAR81695	Analogue	971	5	55.6	4	2	AAW11946	Dimeric p
899	5	55.6	4	2	AAR71564	Hepatitis	972	5	55.6	4	2	AAR74917	Urea plas
900	5	55.6	4	2	AAR90219	Conserved	973	5	55.6	4	2	AAR62426	Accelerat

974 5 55.6 4 2 AAR62433 Accelerat  
975 5 55.6 4 2 AAR67280 Soybean a  
976 5 55.6 4 2 AAR70187 Arg-gingi  
977 5 55.6 4 2 AAR73944 Endotoxin  
978 5 55.6 4 2 AAR73943 Endotoxin  
979 5 55.6 4 2 AAR73942 Endotoxin  
980 5 55.6 4 2 AAR90898 Mu-optoid  
981 5 55.6 4 2 AAR90899 Mu-optoid  
982 5 55.6 4 2 AAR90887 Mu-optoid  
983 5 55.6 4 2 AAR90893 Mu-optoid  
984 5 55.6 4 2 AAR90912 Mu-optoid  
985 5 55.6 4 2 AAR90906 Mu-optoid  
986 5 55.6 4 2 AAR90903 Mu-optoid  
987 5 55.6 4 2 AAR90907 Mu-optoid  
988 5 55.6 4 2 AAR75530 Peptide f  
989 5 55.6 4 2 AAW29563 RGD pepti  
990 5 55.6 4 2 AAR83078 Class I M  
991 5 55.6 4 2 AAR83082 Class I M  
992 5 55.6 4 2 AAR83083 Class I M  
993 5 55.6 4 2 AAR83084 Class I M  
994 5 55.6 4 2 AAR83079 Class I M  
995 5 55.6 4 2 AAR83081 Class I M  
996 5 55.6 4 2 AAR93665 HIV princ  
997 5 55.6 4 2 AAR79634 Endocardi  
998 5 55.6 4 2 AAR66089 Myelopoie  
999 5 55.6 4 2 AAR69077 Leaving u  
1000 5 55.6 4 9 AEB46028 Amino aci

ALIGNMENTS

RESULT 1  
ADH29628  
ID ADH29628 standard; peptide; 2 AA.  
XX AC ADH29628;  
XX 11-MAR-2004 (first entry)  
XX Swinepox homology vector 538-46.26 junction region B peptide.  
DE Swinepox virus; viral vector; homology vector; vaccine; antigen; tumour;  
XX Swinepox virus; viral vector; homology vector; vaccine; antigen; tumour;  
KW cytokine; immune response; feline immunodeficiency virus infection;  
KW heartworm.  
XX Synthetic.  
OS Swinepox virus.  
OS Newcastle disease virus.  
XX WO9622363-A1.  
XX 25-JUL-1996.  
XX 19-JAN-1996; 96WO-US001485.  
XX 19-JAN-1995; 95US-00375992.  
PR 07-JUN-1995; 95US-00472679.  
PR 07-JUN-1995; 95US-00480640.  
PR 07-JUN-1995; 95US-00488237.  
XX (SYTR ) SYNTRO CORP.  
XX Cochran MD, Junker DE;  
PI WPI; 1996-354520/35.  
DR N-PSDB; ADH29627.  
XX Recombinant swine:pox virus contg. foreign DNA sequence - useful for  
PT delivery of vaccinating antigens or other therapeutic agents to humans or  
PT animals.  
XX Disclosure; SEQ ID NO 33; 502pp; English.

XX The invention relates to a new recombinant swinepox virus (SPV)  
CC comprising, inserted into a HindIII M, N or K fragment of the SPV genome,  
CC a foreign DNA sequence that can be expressed in a SPV-infected host cell.  
CC Also new are homology vectors for production of recombinant SPV  
CC comprising double-stranded foreign sequence with, on both sides, double-  
CC stranded SPV DNA homologous to the viral genome on either side of the  
CC HindIII N fragment. The recombinant SPV are vectors for delivering  
CC vaccinating antigens or therapeutic agents to humans, other mammals or  
CC birds. The foreign DNA sequence may encode an antigen from an infectious  
CC agent or tumour, or a cytokine to stimulate an immune response. SPV can  
CC also be used as diagnostic reagents, e.g. to detect feline  
CC immunodeficiency virus of D. immitis (heartworm) infection. SPV is only  
CC weakly pathogenic, species specific and induces an immune response. The  
CC present sequence is encoded by a junction region or fragment from a SPV  
CC homology vector of the invention.  
XX SQ Sequence 2 AA;  
Query Match 55.6%; Score 5; DB 2; Length 2;  
Best Local Similarity 100.0%; Pred. No. 1.9e+06;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 R 1  
Db 2 R 2  
RESULT 2  
AAM98426  
ID AAM98426 standard; peptide; 2 AA.  
XX AC AAM98426;  
XX 24-JAN-2002 (first entry)  
XX Human peptide #1701 encoded by a SNP oligonucleotide.  
DE Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;  
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylose; cancer;  
KW amyloid protein; angiotensin; apoptosis related protein; cadherin;  
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;  
KW complement related protein; cytochrome; kinesin; cytokine; interferon;  
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;  
KW multifactorial disease; autoimmune disease; infection;  
KW nervous system disease.  
XX Homo sapiens.  
OS WO200147944-A2.  
XX 05-JUL-2001.  
XX 28-DEC-2000; 2000WO-US035498.  
XX 28-DEC-1999; 99US-0173419P.  
PR 27-DEC-2000; 2000US-00173419.  
XX (CURA-) CURAGEN CORP.  
XX Shinkets RA, Leach M;  
XX WPI; 2001-465210/50.  
XX Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,  
PT oncogenes and histones, useful for diagnosing and treating, e.g. cancer,  
PT autoimmune diseases and infections.  
XX Disclosure; Page 4041; 4143pp; English.  
PS The present invention relates to oligonucleotides (see AAL26793-AAL34659)  
CC encoding polymorphic variants of proteins related to amylases, amyloid  
CC proteins, angiotensin, apoptosis related proteins, cadherin, cyclin,  
CC

CC polymerase, oncogenes, histones, kinases, colony stimulating factors,  
 CC complement related proteins, cytochromes, kinesins, cytokines,  
 CC interferons, interleukins, G-protein coupled receptors and thioesterases.  
 CC The present sequence is a peptide encoded by one such oligonucleotide.  
 CC The oligonucleotides and the peptides encoded by them may be used in the  
 CC prevention, diagnosis and treatment of diseases associated with  
 CC inappropriate expression of the proteins listed above. Disorders that may  
 CC be prevented, diagnosed and/or treated include multifactorial diseases  
 CC with a genetic component, such as autoimmune diseases (e.g. rheumatoid  
 CC arthritis, multiple sclerosis, diabetes, systemic lupus erythematosus  
 CC and Grave's disease), inflammation, cancer (e.g. cancers of the bladder,  
 CC brain, breast, colon and kidney, leukaemia), diseases of the nervous  
 CC system and an infection of pathogenic organisms  
 XX  
 SQ Sequence 2 AA;

Query Match 55.6%; Score 5; DB 4; Length 2;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+06;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1  
 |  
 Db 1 R 1

RESULT 3  
 AAB91738  
 ID AAB91738 standard; peptide; 2 AA.

AC AAB91738;

DT 22-JUN-2001 (first entry)

DE Opioid peptide SEQ ID NO:914.

XX Protection; endogenous therapeutic peptide; peptidase; conjugation;  
 KW blood component; modification; succinimidy; maleimido group; amino;  
 KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.

XX Homo sapiens.

OS Synthetic.

XX WO200069900-A2.

XX 23-NOV-2000.

XX 17-MAY-2000; 2000WO-US013576.

XX 17-MAY-1999; 99US-0134406P.

PR 10-SEP-1999; 99US-0153406P.

PR 15-OCT-1999; 99US-0159783P.

XX (CONJ-) CONJUCHEM INC.

XX Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudau K;

XX WPI; 2001-112059/12.

XX Modifying and attaching therapeutic peptides to albumin prevents  
 PT peptidase degradation, useful for increasing length of in vivo activity.

XX Disclosure; Page 492; 733pp; English.

XX The present invention describes a modified therapeutic peptide (I)  
 CC comprising a therapeutically active amino acid region (iii) and a  
 CC reactive group (ii) (e.g. succinimidy and maleimido groups) attached to  
 CC a less therapeutically active amino acid region (iv), which covalently  
 CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
 CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
 CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
 CC factors and neurotransmitters, to protect them from peptidase activity in  
 CC vivo for the treatment of various disorders. Endogenous therapeutic  
 CC peptides are not suitable as drug candidates as they require frequent

CC administration due to rapid degradation by peptidases in the body.  
 CC Modifying and attaching therapeutic peptides to albumin prevents or  
 CC reduces the action of peptidases to increase length of activity (half  
 CC life) and specificity as bonding to large molecules decreases  
 CC intracellular uptake and interference with physiological processes.  
 CC AAB90829 to AAB92441 represent peptides which can be used in the  
 CC exemplification of the present invention  
 XX

SQ Sequence 2 AA;

Query Match 55.6%; Score 5; DB 4; Length 2;

Best Local Similarity 100.0%; Pred. No. 1.9e+06;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1  
 |  
 Db 2 R 2

RESULT 4  
 ABG93548

ID ABG93548 standard; peptide; 2 AA.

AC ABG93548;

DT 25-NOV-2002 (first entry)

DE Human P-glycoprotein tryptic peptide #68.

XX Human; P-glycoprotein; tryptic digest; proteolytic cleavage product;  
 KW diabetes; Parkinson's disease; Alzheimer's disease; malaria; cholera;  
 KW human immunodeficiency virus infection; influenza; rabies; diphtheria;  
 KW cancer; multi-drug resistance; MDR.

OS Homo sapiens.

XX EPI223534-A1.

XX 17-JUL-2002.

XX 11-JAN-2002; 2002EP-00075095.

XX 14-JAN-2001; 2001IL-00140881.

PR 19-OCT-2001; 2001US-00982172.

XX (KATZ/) KATZ E I.

XX Katz EI;

XX WPI; 2002-645691/70.

XX Generating amino acid sequences representative of desired polypeptide, by  
 PT computationally generating proteolytic cleavage products, analyzing and  
 PT selecting the set of products, thus generating amino acid sequences.

XX Example 1; Page 14; 124pp; English.

XX The invention relates to generating set of amino acid sequences (AAS)  
 CC representative of one desired polypeptide (I), involving computationally  
 CC generating a number of proteolytic cleavage products (PCP) from (I),  
 CC analysing the PCP according to one parameter defining a characteristic of  
 CC AAS and selecting a set of PCP according to a preset criteria for each  
 CC parameter, thus generating the set of AAS representative of (I). Also  
 CC included are (1) a computer readable storage media (II) comprising a  
 CC database of amino acid sequences corresponding to the polypeptide of  
 CC interest; (2) a system (III) for generating a database of amino acid  
 CC sequences corresponding to a polypeptide of interest, comprises a  
 CC processing unit which executes a software application configured for  
 CC generating the number of proteolytic cleavage products from one  
 CC polypeptide of interest, and analysing the number of proteolytic cleavage  
 CC products according to one parameter defining a characteristic of amino  
 CC acid sequence; (3) a kit for quantifying at least one polypeptide of  
 CC interest, comprises a number of peptides or antibodies each capable of



PT Generating amino acid sequences representative of desired polypeptide, by  
 PT computationally generating proteolytic cleavage products, analyzing and  
 PT selecting the set of products, thus generating amino acid sequences.

PS Example 1; Page 13; 124pp; English.

XX The invention relates to generating set of amino acid sequences (AAS)  
 CC representative of one desired polypeptide (I), involving computationally  
 CC generating a number of proteolytic cleavage products (PCP) from (I),  
 CC analysing the PCP according to one parameter defining a characteristic of  
 CC AAS and selecting a set of PCP according to a preset criteria for each  
 CC parameter, thus generating the set of AAS representative of (I). Also  
 CC included are (1) a computer readable storage media (II) comprising a  
 CC database of amino acid sequences corresponding to the polypeptide of  
 CC interest; (2) a system (III) for generating a database of amino acid  
 CC sequences corresponding to a polypeptide of interest, comprises a  
 CC processing unit which executes a software application configured for  
 CC generating the number of proteolytic cleavage products from one  
 CC polypeptide of interest, and analysing the number of proteolytic cleavage  
 CC products according to one parameter defining a characteristic of amino  
 CC acid sequence; (3) a kit for quantifying at least one polypeptide of  
 CC interest, comprises a number of peptides or antibodies each capable of  
 CC specifically recognising at least one peptide, where the number of  
 CC peptides is generated according to information derived from computational  
 CC analysis of the polypeptide of interest; and (4) quantifying one  
 CC polypeptide of interest in a biological sample, involving contacting the  
 CC biological sample with proteolytic agent, so as to obtain a proteolysed  
 CC biological sample, contacting the proteolysed biological sample with at  
 CC least one antibody and at least one peptide of a number of peptides, and  
 CC detecting presence, absence and/or level of antibody binding to thus  
 CC quantify one polypeptide of interest in the biological sample. The method  
 CC is useful for generating at least one antibody specific to a polypeptide  
 CC of interest. The peptides or antibodies generated may be used to diagnose  
 CC diabetes, Parkinson's disease, Alzheimer's disease, human  
 CC immunodeficiency virus infection, malaria, cholera, influenza, rabies,  
 CC diphtheria, cancer (e.g. breast, colon, cervix, melanoma, lung, ovary,  
 CC pancreas, prostate, lymphomas and leukaemias). The present sequence is a  
 CC predicted tryptic peptide from human P-glycoprotein generated to form  
 CC part of a kit for identifying multi-drug (MDR) resistance associated  
 CC proteins

XX Sequence 2 AA;

Query Match 55.6%; Score 5; DB 5; Length 2;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+06;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1  
 |  
 Db 2 R 2

RESULT 7  
 ABG93615  
 ID ABG93615 standard; peptide; 2 AA.  
 AC ABG93615;  
 XX  
 XX 25-NOV-2002 (first entry)  
 XX Human P-glycoprotein tryptic peptide #135.

XX Human; P-glycoprotein; tryptic digest; proteolytic cleavage product;  
 KW diabetes; Parkinson's disease; Alzheimer's disease; malaria; cholera;  
 KW human immunodeficiency virus infection; influenza; rabies; diphtheria;  
 KW cancer; multi-drug resistance; MDR.

OS Homo sapiens.  
 XX  
 XX EPI223534-A1.  
 PN  
 XX 17-JUL-2002.  
 PD  
 XX

PF 11-JAN-2002; 2002EP-00075095.  
 XX  
 PR 14-JAN-2001; 2001IL-00140881.  
 PR 19-OCT-2001; 2001US-00982172.  
 XX  
 PA (KATZ/) KATZ E I.  
 XX  
 XX Katz EI;

DR WPI; 2002-645691/70.

XX Generating amino acid sequences representative of desired polypeptide, by  
 PT computationally generating proteolytic cleavage products, analyzing and  
 PT selecting the set of products, thus generating amino acid sequences.

XX Example 1; Page 15; 124pp; English.

XX The invention relates to generating set of amino acid sequences (AAS)  
 CC representative of one desired polypeptide (I), involving computationally  
 CC generating a number of proteolytic cleavage products (PCP) from (I),  
 CC analysing the PCP according to one parameter defining a characteristic of  
 CC AAS and selecting a set of PCP according to a preset criteria for each  
 CC parameter, thus generating the set of AAS representative of (I). Also  
 CC included are (1) a computer readable storage media (II) comprising a  
 CC database of amino acid sequences corresponding to the polypeptide of  
 CC interest; (2) a system (III) for generating a database of amino acid  
 CC sequences corresponding to a polypeptide of interest, comprises a  
 CC processing unit which executes a software application configured for  
 CC generating the number of proteolytic cleavage products from one  
 CC polypeptide of interest, and analysing the number of proteolytic cleavage  
 CC products according to one parameter defining a characteristic of amino  
 CC acid sequence; (3) a kit for quantifying at least one polypeptide of  
 CC interest, comprises a number of peptides or antibodies each capable of  
 CC specifically recognising at least one peptide, where the number of  
 CC peptides is generated according to information derived from computational  
 CC analysis of the polypeptide of interest; and (4) quantifying one  
 CC polypeptide of interest in a biological sample, involving contacting the  
 CC biological sample with proteolytic agent, so as to obtain a proteolysed  
 CC biological sample, contacting the proteolysed biological sample with at  
 CC least one antibody and at least one peptide of a number of peptides, and  
 CC detecting presence, absence and/or level of antibody binding to thus  
 CC quantify one polypeptide of interest in the biological sample. The method  
 CC is useful for generating at least one antibody specific to a polypeptide  
 CC of interest. The peptides or antibodies generated may be used to diagnose  
 CC diabetes, Parkinson's disease, Alzheimer's disease, human  
 CC immunodeficiency virus infection, malaria, cholera, influenza, rabies,  
 CC diphtheria, cancer (e.g. breast, colon, cervix, melanoma, lung, ovary,  
 CC pancreas, prostate, lymphomas and leukaemias). The present sequence is a  
 CC predicted tryptic peptide from human P-glycoprotein generated to form  
 CC part of a kit for identifying multi-drug (MDR) resistance associated  
 CC proteins

XX Sequence 2 AA;

Query Match 55.6%; Score 5; DB 5; Length 2;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+06;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1  
 |  
 Db 2 R 2

RESULT 8  
 ABG93621  
 ID ABG93621 standard; peptide; 2 AA.  
 AC ABG93621;  
 XX  
 XX 25-NOV-2002 (first entry)

XX Human P-glycoprotein tryptic peptide #141.  
 DE

KW Human; P-glycoprotein; tryptic digest; proteolytic cleavage product;  
 KW diabetes; Parkinson's disease; Alzheimer's disease; malaria; cholera;  
 KW human immunodeficiency virus infection; influenza; rabies; diphtheria;  
 KW cancer; multi-drug resistance; MDR.  
 XX Homo sapiens.  
 XX EPI223534-A1.  
 XX 17-JUL-2002.  
 XX 11-JAN-2002; 2002EP-00075095.  
 XX 14-JAN-2001; 2001IL-00140881.  
 XX 19-OCT-2001; 2001US-00982172.  
 XX (KATZ/) KATZ E I.  
 XX Katz EI;  
 XX WPI; 2002-645691/70.  
 XX Generating amino acid sequences representative of desired polypeptide, by  
 XX computationally generating proteolytic cleavage products, analyzing and  
 XX selecting the set of products, thus generating amino acid sequences.  
 XX Example 1; Page 15; 124pp; English.  
 XX The invention relates to generating set of amino acid sequences (AAS)  
 XX representative of one desired polypeptide (I), involving computationally  
 XX generating a number of proteolytic cleavage products (PCP) from (I),  
 XX analysing the PCP according to one parameter defining a characteristic of  
 XX AAS and selecting a set of PCP according to a preset criteria for each  
 XX parameter, thus generating the set of AAS representative of (I). Also  
 XX included are (1) a computer readable storage media (II) comprising a  
 XX database of amino acid sequences corresponding to the polypeptide of  
 XX interest; (2) a system (III) for generating a database of amino acid  
 XX sequences corresponding to a polypeptide of interest, comprises a  
 XX processing unit which executes a software application configured for  
 XX generating the number of proteolytic cleavage products from one  
 XX polypeptide of interest, and analysing the number of proteolytic cleavage  
 XX products according to one parameter defining a characteristic of amino  
 XX acid sequence; (3) a kit for quantifying at least one polypeptide of  
 XX interest, comprises a number of peptides or antibodies each capable of  
 XX specifically recognising at least one peptide, where the number of  
 XX peptides is generated according to information derived from computational  
 XX analysis of the polypeptide of interest; and (4) quantifying one  
 XX polypeptide of interest in a biological sample, involving contacting the  
 XX biological sample with proteolytic agent, so as to obtain a proteolysed  
 XX sample, contacting the proteolysed biological sample with at  
 XX least one antibody and at least one peptide of a number of peptides, and  
 XX detecting presence, absence and/or level of antibody binding to thus  
 XX quantify one polypeptide of interest in the biological sample. The method  
 XX is useful for generating at least one antibody specific to a polypeptide  
 XX of interest. The peptides or antibodies generated may be used to diagnose  
 XX diabetes, Parkinson's disease, Alzheimer's disease, human  
 XX immunodeficiency virus infection, malaria, cholera, influenza, rabies,  
 XX diphtheria, cancer (e.g. breast, colon, cervix, melanoma, lung, ovary,  
 XX pancreas, prostate, lymphomas and leukaemias). The present sequence is a  
 XX predicted tryptic peptide from human P-glycoprotein generated to form  
 XX part of a kit for identifying multi-drug (MDR) resistance associated  
 XX proteins  
 XX Sequence 2 AA;

Query Match 55.6%; Score 5; DB 5; Length 2;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+06;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1  
 2 R 2  
 Db

RESULT 9  
 ID ABG93576 standard; peptide; 2 AA.  
 XX AC ABG93576;  
 XX 25-NOV-2002 (first entry)  
 XX Human P-glycoprotein tryptic peptide #96.  
 XX Human; P-glycoprotein; tryptic digest; proteolytic cleavage product;  
 KW diabetes; Parkinson's disease; Alzheimer's disease; malaria; cholera;  
 KW human immunodeficiency virus infection; influenza; rabies; diphtheria;  
 KW cancer; multi-drug resistance; MDR.  
 XX Homo sapiens.  
 XX EPI223534-A1.  
 XX 17-JUL-2002.  
 XX 11-JAN-2002; 2002EP-00075095.  
 XX 14-JAN-2001; 2001IL-00140881.  
 XX 19-OCT-2001; 2001US-00982172.  
 XX (KATZ/) KATZ E I.  
 XX Katz EI;  
 XX WPI; 2002-645691/70.  
 XX Generating amino acid sequences representative of desired polypeptide, by  
 XX computationally generating proteolytic cleavage products, analyzing and  
 XX selecting the set of products, thus generating amino acid sequences.  
 XX Example 1; Page 14; 124pp; English.  
 XX The invention relates to generating set of amino acid sequences (AAS)  
 XX representative of one desired polypeptide (I), involving computationally  
 XX generating a number of proteolytic cleavage products (PCP) from (I),  
 XX analysing the PCP according to one parameter defining a characteristic of  
 XX AAS and selecting a set of PCP according to a preset criteria for each  
 XX parameter, thus generating the set of AAS representative of (I). Also  
 XX included are (1) a computer readable storage media (II) comprising a  
 XX database of amino acid sequences corresponding to the polypeptide of  
 XX interest; (2) a system (III) for generating a database of amino acid  
 XX sequences corresponding to a polypeptide of interest, comprises a  
 XX processing unit which executes a software application configured for  
 XX generating the number of proteolytic cleavage products from one  
 XX polypeptide of interest, and analysing the number of proteolytic cleavage  
 XX products according to one parameter defining a characteristic of amino  
 XX acid sequence; (3) a kit for quantifying at least one polypeptide of  
 XX interest, comprises a number of peptides or antibodies each capable of  
 XX specifically recognising at least one peptide, where the number of  
 XX peptides is generated according to information derived from computational  
 XX analysis of the polypeptide of interest; and (4) quantifying one  
 XX polypeptide of interest in a biological sample, involving contacting the  
 XX biological sample with proteolytic agent, so as to obtain a proteolysed  
 XX sample, contacting the proteolysed biological sample with at  
 XX least one antibody and at least one peptide of a number of peptides, and  
 XX detecting presence, absence and/or level of antibody binding to thus  
 XX quantify one polypeptide of interest in the biological sample. The method  
 XX is useful for generating at least one antibody specific to a polypeptide  
 XX of interest. The peptides or antibodies generated may be used to diagnose  
 XX diabetes, Parkinson's disease, Alzheimer's disease, human  
 XX immunodeficiency virus infection, malaria, cholera, influenza, rabies,  
 XX diphtheria, cancer (e.g. breast, colon, cervix, melanoma, lung, ovary,  
 XX pancreas, prostate, lymphomas and leukaemias). The present sequence is a  
 XX predicted tryptic peptide from human P-glycoprotein generated to form  
 XX part of a kit for identifying multi-drug (MDR) resistance associated  
 XX proteins





Qy 1 R 1  
Db 1 R 1

RESULT 12  
ABR39485  
ID ABR39485 standard; peptide; 2 AA.  
XX ABR39485;  
XX  
DT 12-JUN-2003 (first entry)  
XX  
XX AlphaS1 casein N-terminal fragment derived peptide #74.  
XX  
XX AlphaS1 casein; haematopoiesis; virucide; immunosuppressive; antilipemic;  
KW immunoestimulant; antidiabetic; anti-HIV; cytostatic; antibacterial;  
KW antianemic.  
XX  
XX Synthetic.  
OS Bos sp.  
XX  
XX WO2003018606-A2.  
FN  
XX  
XX 06-MAR-2003.  
PD  
XX  
XX 29-AUG-2002; 2002WO-IL000720.  
PF  
XX  
XX 30-AUG-2001; 2001US-00942121.  
PR  
XX  
XX (CHAY-) CHAY 13 MEDICAL RES GROUP NV.  
PA  
XX  
XX Sidelman Z;  
PI  
XX  
XX WPI; 2003-312868/30.  
DR  
XX  
XX New purified peptide sequences, useful for treating e.g. autoimmune  
PT disorders, derived by fragmentation of an N terminus of alpha-S1 casein  
PT or synthesis.  
PT  
XX  
XX Claim 4; Page 78; 190pp; English.  
PS  
XX  
XX Sequences ABR39485-509 represent synthetic peptides derived from the N-  
CC terminal fragment (2-26 residues) of alphaS1 casein. The peptides  
CC stimulate and enhance immune response, protect against viral infection,  
CC normalize serum cholesterol level, stimulates haematopoiesis and is non-  
CC toxic. They are useful for treating a viral disease or infection, an  
CC autoimmune disease, haematopoiesis; for inducing haematopoietic stem  
CC cells proliferation and differentiation, megakaryocytopoiesis,  
CC erythropoiesis, leukocytopoiesis, thrombocytopoiesis, plasma cell  
CC proliferation, dendritic cell proliferation, macrophage proliferation;  
CC for preventing thrombocytopenia, pancytopenia, granulocytopenia,  
CC hyperlipidemia, cholesteraemia, glucosuria, diabetes, AIDS infection by  
CC supported by autologous bone marrow or peripheral blood stem cell  
CC transplantation (ASCT) or allogeneic bone marrow transplantation (BMT),  
CC erythropoietin treatable conditions; for augmenting the effect of  
CC erythropoietin or thrombopoietin; for treating thrombopoietin treatable  
CC condition; for enhancing peripheral stem cell mobilization; for treating  
CC haematological disease and deficiencies, hypercholesterolemia,  
CC hyperglycemia, helper T-cell disorders, dendritic cell deficiencies,  
CC haematopoietic stem cell disorders (including platelet, lymphocyte,  
CC plasma cell and neutrophil disorders) pre-leukemic conditions, leukemic  
CC conditions, immune system disorders resulting from chemotherapy or  
CC radiation therapy, human immune system disorders resulting from treatment  
CC of diseases of immune deficiency and bacterial infections,  
CC myelodysplastic syndrome, aplastic anemia and bone marrow insufficiency;  
CC for enhancing colonization of blood stem cells in a myeloablated  
XX recipient  
XX  
SQ Sequence 2 AA;

Query Match 55.6%; Score 5; DB 6; Length 2;  
Best Local Similarity 100.0%; Pred. No. 1.9e+06;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1  
Db 1 R 1

RESULT 13  
AAE39209  
ID AAE39209 standard; peptide; 2 AA.  
XX  
XX AAE39209;  
AC  
XX  
XX 18-DEC-2003 (first entry)  
DT  
XX  
XX Angiogenesis inhibitor peptide motif #38.  
DE  
XX  
XX Angiogenesis inhibitor; methionine aminopeptidase-2; MetAP-2; trachoma;  
KW autoimmune disease; rheumatoid arthritis; retinopathy of prematurity;  
KW ocular disease; diabetic retinopathy; corneal graft rejection; cancer;  
KW retrolental fibroplasia; neovascular glaucoma; rubeosis; ocular tumour;  
KW immunosuppressive; cytostatic; ophthalmological.  
KW  
XX  
XX Unidentified.  
OS  
XX  
XX Key Location/Qualifiers  
FH Modified-site 2 /note= "Arg(3-amino 3-pyriylpropionic acid)"  
FT  
XX  
XX US2002193298-A1.  
FN  
XX  
XX 19-DEC-2002.  
PD  
XX  
XX 05-OCT-2001; 2001US-00972772.  
PF  
XX  
XX 01-NOV-2000; 2000US-00704251.  
PR  
XX  
XX (PRAE-) PRAECIS PHARM INC.  
PA  
XX  
XX Olson GL, Self C, Lee L, Cook CM, Birktoft J;  
PI  
XX  
XX WPI; 2003-755034/71.  
DR  
XX  
XX New angiogenesis inhibitor compound useful for treating angiogenic  
PT diseases such as cancer, diabetic retinopathy, hypoxia, ocular tumors,  
PT trachoma.  
PT  
XX  
XX Disclosure; Page 5; 38pp; English.  
PS  
XX  
XX The present invention provides novel angiogenesis inhibitor compounds  
CC comprising a methionine aminopeptidase (MetAP)-2 inhibitory core coupled  
CC to a peptide. The invention is useful for treating an angiogenic disease  
CC such as autoimmune disease such as rheumatoid arthritis and cancer. The  
CC invention is also useful for treating angiogenic diseases such as ocular  
CC disease e.g. diabetic retinopathy, retinopathy of prematurity, corneal  
CC graft rejection, retrolental fibroplasia, neovascular glaucoma, rubeosis,  
CC angiogenesis in eye associated with infection or surgical intervention,  
CC ocular tumours and trachoma. The present sequence is angiogenesis  
CC inhibitor peptide motif  
XX  
SQ Sequence 2 AA;

Query Match 55.6%; Score 5; DB 7; Length 2;  
Best Local Similarity 100.0%; Pred. No. 1.9e+06;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1  
Db 2 R 2

```

RESULT 14
ADL98393
ID ADL98393 standard; peptide; 2 AA.
XX
AC ADL98393;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human leukocyte antigen-B15 allotype B*1508/B*1501 peptide #27.
XX
KW soluble human leukocyte antigen; HLA; locus-specific primer; truncation;
KW multimeric HLA complex; bioreactor;
KW major histocompatibility complex molecule; MHC; vaccine; HLA-B15;
KW allotype; B*1508; B*1501.
XX
OS Homo sapiens.
XX
PN US2003166057-A1.
XX
PD 04-SEP-2003.
XX
PF 18-DEC-2001; 2001US-00022066.
XX
PR 17-DEC-1999; 99US-00465321.
PR 18-DEC-2000; 2000US-0256409P.
PR 18-DEC-2000; 2000US-0256410P.
PR 24-MAY-2001; 2001US-0293261P.
PR 09-OCT-2001; 2001US-0327907P.
PR 10-OCT-2001; 2001US-00974366.
XX
PA (HILD/) HILDEBRAND W H.
PA (PRIL/) PRILLIMAN K R.
XX
FI Hildebrand WH, Prilliman KR;
XX
DR WPI; 2003-863700/80.
XX
PT Producing soluble human leukocyte antigen molecules, for testing the
PT functionality of peptide ligands, comprises utilizing a locus-specific
PT primer having a stop codon incorporated into a 3' primer, or that
PT truncates the allelic cDNA.
XX
PS Disclosure; Fig 26; 148pp; English.
XX
CC The invention describes a method of producing soluble human leukocyte
CC antigen (HLA) molecules comprising utilizing a locus-specific primer
CC having a stop codon incorporated into a 3' primer, or a locus-specific
CC primer that truncates the allelic cDNA resulting in a truncated PCR
CC product having the coding regions encoding cytoplasmic and transmembrane
CC domains of the allelic cDNA removed so that the truncated PCR product has
CC a coding region encoding a soluble HLA molecule. Also described is a
CC multimeric HLA complex comprising a substrate, and at least two soluble
CC HLA molecules attached to the substrate and an apparatus or a bioreactor
CC unit for producing major histocompatibility complex molecules. The
CC methods are useful for producing soluble human leukocyte antigen (HLA)
CC molecules. The multimeric HLA complex is useful for testing the
CC functionality of peptide ligands bound to the soluble HLA molecules. The
CC HLA molecules are also useful in vaccine development. This is the amino
CC acid sequence of a human leukocyte antigen (HLA)-B15 peptide found in
CC allotypes B*1508 and B*1501.
XX
SQ Sequence 2 AA;
Query Match 55.6%; Score 5; DB 7; Length 2;
Best Local Similarity 100.0%; Pred. No. 1.9e+06;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 R 1
Db 2 R 2
RESULT 15
ADL98393
ID ADL98393 standard; peptide; 2 AA.
XX
AC ADL98393;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human leukocyte antigen-B15 allotype B*1508/B*1501 peptide #27.
XX
KW soluble human leukocyte antigen; HLA; locus-specific primer; truncation;
KW multimeric HLA complex; bioreactor;
KW major histocompatibility complex molecule; MHC; vaccine; HLA-B15;
KW allotype; B*1508; B*1501.
XX
OS Homo sapiens.
XX
PN US2003166057-A1.
XX
PD 04-SEP-2003.
XX
PF 18-DEC-2001; 2001US-00022066.
XX
PR 17-DEC-1999; 99US-00465321.
PR 18-DEC-2000; 2000US-0256409P.
PR 18-DEC-2000; 2000US-0256410P.
PR 24-MAY-2001; 2001US-0293261P.
PR 09-OCT-2001; 2001US-0327907P.
PR 10-OCT-2001; 2001US-00974366.
XX
PA (HILD/) HILDEBRAND W H.
PA (PRIL/) PRILLIMAN K R.
XX
FI Hildebrand WH, Prilliman KR;
XX
DR WPI; 2003-863700/80.
XX
PT Producing soluble human leukocyte antigen molecules, for testing the
PT functionality of peptide ligands, comprises utilizing a locus-specific
PT primer having a stop codon incorporated into a 3' primer, or that
PT truncates the allelic cDNA.
XX
PS Disclosure; Fig 26; 148pp; English.
XX
CC The invention describes a method of producing soluble human leukocyte
CC antigen (HLA) molecules comprising utilizing a locus-specific primer
CC having a stop codon incorporated into a 3' primer, or a locus-specific
CC primer that truncates the allelic cDNA resulting in a truncated PCR
CC product having the coding regions encoding cytoplasmic and transmembrane
CC domains of the allelic cDNA removed so that the truncated PCR product has
CC a coding region encoding a soluble HLA molecule. Also described is a
CC multimeric HLA complex comprising a substrate, and at least two soluble
CC HLA molecules attached to the substrate and an apparatus or a bioreactor
CC unit for producing major histocompatibility complex molecules. The
CC methods are useful for producing soluble human leukocyte antigen (HLA)
CC molecules. The multimeric HLA complex is useful for testing the
CC functionality of peptide ligands bound to the soluble HLA molecules. The
CC HLA molecules are also useful in vaccine development. This is the amino
CC acid sequence of a human leukocyte antigen (HLA)-B15 peptide found in
CC allotypes B*1508 and B*1501.
XX
SQ Sequence 2 AA;
Query Match 55.6%; Score 5; DB 7; Length 2;
Best Local Similarity 100.0%; Pred. No. 1.9e+06;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 R 1
Db 2 R 2
RESULT 16
ADW36964
ID ADW36964 standard; peptide; 2 AA.
XX
AC ADW36964;
XX
DT 10-MAR-2005 (first entry)
XX
DE HLA binding epitope #7391.
XX
KW Virucide; cytostatic; gene therapy; vaccine; epitope; cytotoxic T cell;
KW MHC class I; CTL; HTL; A2-restricted cytotoxic lymphocyte; HLA;
KW viral disease; cancer.
XX
OS Unidentified.
XX
PN WO2003040165-A2.
XX
PD 15-MAY-2003.
XX
PF 18-OCT-2001; 2001WO-US051650.
XX
PR 19-OCT-2000; 2000US-0242350P.
PR 20-APR-2001; 2001US-0285624P.
XX
PA (EPIM-) EPIMUNE INC.
XX
PI Sette A, Sidney J, Southwood S;
XX
DR WPI; 2003-441519/41.
XX
PT New composition comprising at least one peptide having allele-specific
PT binding motifs for HLA, useful for preventing, treating or diagnosing
PT viral diseases and cancer.
XX
PS Claim 1; Page 52-379; 382pp; English.
XX
CC The invention relates to a composition comprising at least one peptide
CC having an isolated, prepared epitope selected from any of the sequences
CC from 30 lists given in the specification. Also disclosed is a method for
CC inducing a cytotoxic T cell response against a pre-selected antigen in a
CC patient expressing a specific MHC class I allele by contacting cytotoxic
CC T cells from the patient with the composition cited above. The
CC composition comprises an epitope that is joined by an amino acid linker.
CC The epitope is admixed or joined to a CTL or HTL epitope. The epitope is
CC bound to an HLA molecule on the antigen-presenting cell, where when an A2
CC -restricted cytotoxic lymphocyte (CTL) is present, a receptor of the CTL
CC binds to a complex of the HLA molecule and the epitope. Specifically
CC claimed are peptides having allele-specific binding motifs for HLA. The
CC compositions and methods are useful for preventing, treating or
CC diagnosing viral diseases and cancer. The peptide epitopes are useful as
CC diagnostic agents for evaluating immune responses, for making antibodies
CC and for evaluating efficacy of a vaccine. Sequences given in ADW29251-
CC ADW37745 represent epitopes of the invention as given in Tables 2-31.
XX
SQ Sequence 2 AA;
Query Match 55.6%; Score 5; DB 7; Length 2;
Best Local Similarity 100.0%; Pred. No. 1.9e+06;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 R 1
Db 2 R 2

```

DE HLA binding epitope #7714.  
 XX Virucide; cytostatic; gene therapy; vaccine; epitope; cytotoxic T cell;  
 KW MHC class I; CTL; HTL; A2-restricted cytotoxic lymphocyte; HLA;  
 KW viral disease; cancer.  
 XX Unidentified.  
 OS  
 XX WO2003040165-A2.  
 FN  
 XX 15-MAY-2003.  
 PD  
 XX 18-OCT-2001; 2001WO-US051650.  
 XX  
 XX 19-OCT-2000; 2000US-0242350P.  
 XX 20-APR-2001; 2001US-0285624P.  
 PR  
 XX (EPIM-) EPIMMUNE INC.  
 PA  
 XX Sette A, Sidney J, Southwood S;  
 PI WPI; 2003-441519/41.  
 DR  
 XX New composition comprising at least one peptide having allele-specific  
 PT binding motifs for HLA, useful for preventing, treating or diagnosing  
 PT viral diseases and cancer.  
 XX  
 PS Claim 1; Page 52-379; 382pp; English.  
 XX  
 XX The invention relates to a composition comprising at least one peptide  
 CC having an isolated, prepared epitope selected from any of the sequences  
 CC from 30 lists given in the specification. Also disclosed is a method for  
 CC inducing a cytotoxic T cell response against a pre-selected antigen in a  
 CC patient expressing a specific MHC class I allele by contacting cytotoxic  
 CC T cells from the patient with the composition cited above. The  
 CC composition comprises an epitope that is joined by an amino acid linker.  
 CC The epitope is admixed or joined to a CTL or HTL epitope. The epitope is  
 CC bound to an HLA molecule on the antigen-presenting cell, where when an A2  
 CC -restricted cytotoxic lymphocyte (CTL) is present, a receptor of the CTL  
 CC binds to a complex of the HLA molecule and the epitope. Specifically  
 CC claimed are peptides having allele-specific binding motifs for HLA. The  
 CC compositions and methods are useful for preventing, treating or  
 CC diagnosing viral diseases and cancer. The peptide epitopes are useful as  
 CC diagnostic agents for evaluating immune responses, for making antibodies  
 CC and for evaluating efficacy of a vaccine. Sequences given in ADW29251-  
 CC ADW37745 represent epitopes of the invention as given in Tables 2-31.  
 XX  
 SQ Sequence 2 AA;  
 Query Match 55.6%; Score 5; DB 7; Length 2;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+06;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 R 1  
 Db 1 R 1  
 RESULT 17  
 ADW37285  
 ID ADW37285 standard; peptide; 2 AA.  
 XX  
 AC ADW37285;  
 XX  
 DT 10-MAR-2005 (first entry)  
 XX  
 DE HLA binding epitope #8035.  
 XX  
 KW Virucide; cytostatic; gene therapy; vaccine; epitope; cytotoxic T cell;  
 KW MHC class I; CTL; HTL; A2-restricted cytotoxic lymphocyte; HLA;  
 KW viral disease; cancer.  
 XX Unidentified.  
 OS

XX WO2003040165-A2.  
 PN  
 XX 15-MAY-2003.  
 PD  
 XX 18-OCT-2001; 2001WO-US051650.  
 XX  
 XX 19-OCT-2000; 2000US-0242350P.  
 PR 20-APR-2001; 2001US-0285624P.  
 PR  
 XX (EPIM-) EPIMMUNE INC.  
 PA  
 XX Sette A, Sidney J, Southwood S;  
 PI WPI; 2003-441519/41.  
 DR  
 XX New composition comprising at least one peptide having allele-specific  
 PT binding motifs for HLA, useful for preventing, treating or diagnosing  
 PT viral diseases and cancer.  
 XX  
 PS Claim 1; Page 52-379; 382pp; English.  
 XX  
 XX The invention relates to a composition comprising at least one peptide  
 CC having an isolated, prepared epitope selected from any of the sequences  
 CC from 30 lists given in the specification. Also disclosed is a method for  
 CC inducing a cytotoxic T cell response against a pre-selected antigen in a  
 CC patient expressing a specific MHC class I allele by contacting cytotoxic  
 CC T cells from the patient with the composition cited above. The  
 CC composition comprises an epitope that is joined by an amino acid linker.  
 CC The epitope is admixed or joined to a CTL or HTL epitope. The epitope is  
 CC bound to an HLA molecule on the antigen-presenting cell, where when an A2  
 CC -restricted cytotoxic lymphocyte (CTL) is present, a receptor of the CTL  
 CC binds to a complex of the HLA molecule and the epitope. Specifically  
 CC claimed are peptides having allele-specific binding motifs for HLA. The  
 CC compositions and methods are useful for preventing, treating or  
 CC diagnosing viral diseases and cancer. The peptide epitopes are useful as  
 CC diagnostic agents for evaluating immune responses, for making antibodies  
 CC and for evaluating efficacy of a vaccine. Sequences given in ADW29251-  
 CC ADW37745 represent epitopes of the invention as given in Tables 2-31.  
 XX  
 SQ Sequence 2 AA;  
 Query Match 55.6%; Score 5; DB 7; Length 2;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+06;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 R 1  
 Db 2 R 2  
 RESULT 18  
 AAP30601  
 ID AAP30601 standard; protein; 3 AA.  
 XX  
 AC AAP30601;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 31-MAY-1992 (first entry)  
 XX  
 DE Sequence of beta-lactamase encoded on pULB1523.  
 XX  
 KW Emphysema therapy; proteolytic enzyme; lung disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN BE895961-A.  
 XX  
 PD 16-JUN-1983.  
 XX  
 PF 21-FEB-1983; 83BE-00210157.  
 XX  
 PR 21-FEB-1983; 83BE-00210157.

```

PR 21-FEB-1983; 83BE-00895961.
XX (REGI-) REGION WALLONNE.
PA (EJEC-) EJECUTIVO REG VALON.
XX
XX WPI; 1983-700089/27.
DR N-PSDB; AAN30207.
XX
PT Bacterial clone producing alpha-1-antitrypsin - transformed with vector
PT contg. antitrypsin DNA.
XX
XX Disclosure; Fig 2; 23pp; French.
XX
CC The inventors claim double-stranded cDNA encoding AT and E.coli clones
CC contg. AT-DNA. The cDNA is derived from mRNA extracted from human liver.
CC Recombinant alpha-1-AT can be used for the treatment of pollutant-induced
CC lung damage (esp. emphysema), particularly when caused by excessive
CC release of proteolytic enzymes. Large amts. of AT can be made by
CC cultivating the transformed bacteria. (Updated on 25-MAR-2003 to correct
CC PR field.)
XX
SQ Sequence 3 AA;

Query Match 55.6%; Score 5; DB 1; Length 3;
Best Local Similarity 100.0%; Pred. No. 1.9e+06;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1
Db 1 R 1

RESULT 19
AAP90668
ID AAP90668 standard; protein; 3 AA.
AC AAP90668;
XX
XX 10-MAR-2003 (revised)
DT 26-MAY-1990 (first entry)
XX
XX New antihypertensive peptide.
DE Antihypertensive peptide; angiotensin-converting enzyme; ACE.
XX
XX Synthetic.
OS
XX JP01083096-A.
FN
XX 28-MAR-1989.
PD
XX 25-SEP-1987; 87JP-00241646.
PF
XX 25-SEP-1987; 87JP-00241646.
PR
XX (AJIN ) AJINOMOTO KK.
PA
XX WPI; 1989-136272/18.
DR
XX New antihypertensive peptide(s) - used to inhibit angiotensin-converting
PT enzyme.
XX
XX Disclosure; Page; 20pp; Japanese.
XX
XX The peptide and its salts inhibit angiotensin-converting enzyme (ACE) and
XX are useful as antihypertensives. They may be administered orally,
XX parenterally or rectally in the form of tablets, capsules, granules,
XX powder, syrup, suspension, suppositories, ointment, cream, gel, plaster,
XX inhalation compsn. or injection at a dose of 0.001-1000, pref. 0.01-10,
XX mg, 1-3 times per day. (Updated on 10-MAR-2003 to add missing OS field.)
XX
XX Sequence 3 AA;

Query Match 55.6%; Score 5; DB 1; Length 3;
Best Local Similarity 100.0%; Pred. No. 1.9e+06;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1
Db 1 R 1

RESULT 20
AAP90665
ID AAP90665 standard; protein; 3 AA.
XX
XX AAP90665;
AC
XX 10-MAR-2003 (revised)
DT 26-MAY-1990 (first entry)
XX
XX New antihypertensive peptide.
DE Antihypertensive peptide; angiotensin-converting enzyme; ACE.
XX
XX Synthetic.
OS
XX JP01083096-A.
FN
XX 28-MAR-1989.
PD
XX 25-SEP-1987; 87JP-00241646.
PF
XX 25-SEP-1987; 87JP-00241646.
PR
XX (AJIN ) AJINOMOTO KK.
PA
XX WPI; 1989-136272/18.
DR
XX New antihypertensive peptide(s) - used to inhibit angiotensin-converting
PT enzyme.
XX
XX Disclosure; Page; 20pp; Japanese.
XX
XX The peptide and its salts inhibit angiotensin-converting enzyme (ACE) and
XX are useful as antihypertensives. They may be administered orally,
XX parenterally or rectally in the form of tablets, capsules, granules,
XX powder, syrup, suspension, suppositories, ointment, cream, gel, plaster,
XX inhalation compsn. or injection at a dose of 0.001-1000, pref. 0.01-10,
XX mg, 1-3 times per day. (Updated on 10-MAR-2003 to add missing OS field.)
XX
XX Sequence 3 AA;

Query Match 55.6%; Score 5; DB 1; Length 3;
Best Local Similarity 100.0%; Pred. No. 1.9e+06;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1
Db 2 R 2

RESULT 21
AAP97811
ID AAP97811 standard; protein; 3 AA.
XX
XX AAP97811;
AC
XX 29-JUL-1992 (first entry)
DT
XX Sequence of fragment 24, the tryptic fragment of recombinant penicillin
DE acyltransferase (PA) polypeptide 2.
XX
XX Penicillin biosynthesis; enzyme; antibiotic.
KW
XX Penicillium chrysogenum.
OS

```

```

XX EP336446-A.
XX
XX 11-OCT-1989.
XX
XX 07-APR-1989; 89EP-00106214.
XX
XX 08-APR-1988; 89AT-00000922.
XX
XX 13-JUL-1988; 89AT-00001806.
XX
XX 08-SEP-1988; 89AT-00002201.
XX
XX (BIOC ) BIOCHEMIE GMBH.
XX
XX Knauseder F, Leitner E, Palma N, Weber G;
XX
XX WPI; 1989-294357/41.
XX
XX Recombinant penicillin acyl-transferase - and DNA coding for it.
XX
XX Claim 9; Page 48; 52pp; English.
XX
XX The inventors claim recombinant penicillin acyltransferase (PAT) and DNA
XX coding for PAT. PAT catalyses the last step in the biosynthesis of
XX penicillin G and penicillin V. More specifically, the coding strand of
XX the DNA has the nucleotide sequence shown below. This includes three
XX introns and codes for a PAT protein with mol. wt. ca. 40KD. Plasmid
XX vectors pBC2001 and pBC2002 are specifically claimed
XX
XX Sequence 3 AA;
XX
XX Query Match 55.6%; Score 5; DB 1; Length 3;
XX Best Local Similarity 100.0%; Pred. No. 1.9e+06;
XX Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 R 1
XX |
XX 3 R 3
XX
XX RESULT 22
XX AAR00718
XX ID AAR00718 standard; peptide; 3 AA.
XX
XX AC AAR00718;
XX
XX DT 09-JAN-2003 (revised)
XX
XX DT 29-MAY-1990 (first entry)
XX
XX DE Core repeat of cell-adhesive protein.
XX
XX KW Cell adhesion; anti-metastatic agent; immunomodulation; core repeat.
XX
XX OS Homo sapiens.
XX
XX XX EP347931-A.
XX
XX PD 27-DEC-1989.
XX
XX PF 23-JUN-1989; 89EP-00111468.
XX
XX PR 24-JUN-1988; 88JP-00156133.
XX
XX (AZUM/) AZUMA I.
XX
XX Saiki I, Nishi N, Azuma I, Tokura S;
XX
XX WPI; 1990-001305/01.
XX
XX Polypeptide with repeated sequences of cell-adhesive protein - used as
XX anti-metastatic agent for cancer and agonist or antagonist of cell-
XX adhesive proteins.
XX
XX Claim 2; Page 14; 16pp; English.
XX
XX
XX Peptide core is repeated 2-20 times to form a cell-adhesive protein of
XX mol. wt. 1,500-5,000. The protein is an (ant)agonist of cell-adhesive
XX proteins such as fibronectin. It has high antimetastatic activity against
XX cancer and can be used in immunomodulation, wound healing, platelet
XX aggregation inhibition and alleviation of neuro-disorders. See also
XX AAR00722. (Updated on 09-JAN-2003 to add missing OS field.)
XX
XX Sequence 3 AA;
XX
XX Query Match 55.6%; Score 5; DB 2; Length 3;
XX Best Local Similarity 100.0%; Pred. No. 1.9e+06;
XX Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 R 1
XX |
XX 1 R 1
XX
XX RESULT 23
XX AAR04607
XX ID AAR04607 standard; peptide; 3 AA.
XX
XX AC AAR04607;
XX
XX DT 25-MAR-2003 (revised)
XX
XX DT 05-SEP-1990 (first entry)
XX
XX DE Antiviral agent.
XX
XX KW Antiviral; M2; poliovirus; polio; hepatitis.
XX
XX OS Synthetic.
XX
XX PN JP02078631-A.
XX
XX PD 19-MAR-1990.
XX
XX PF 14-SEP-1988; 88JP-00228843.
XX
XX PR 14-SEP-1988; 88JP-00228843.
XX
XX (NIHA ) NIPPON MINING CO.
XX
XX WPI; 1990-129060/17.
XX
XX Antiviral agent contg. tripeptide (unit) - of basic aminoacid, then
XX alanine, glycine or sarcosine, and acidic aminoacid, effective against
XX virus with protein-terminated DNA or RNA.
XX
XX Disclosure; Page ?; 4pp; Japanese.
XX
XX Peptide is effective against inhibiting propagation of DNA or RNA bonded,
XX protein containing viruses eg. Poliovirus, Hepatitis virus. (Updated on
XX 25-MAR-2003 to correct PA field.)
XX
XX Sequence 3 AA;
XX
XX Query Match 55.6%; Score 5; DB 2; Length 3;
XX Best Local Similarity 100.0%; Pred. No. 1.9e+06;
XX Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 R 1
XX |
XX 1 R 1
XX
XX RESULT 24
XX AAR10543
XX ID AAR10543 standard; protein; 3 AA.
XX
XX AC AAR10543;
XX
XX

```

DT 25-MAR-2003 (revised)  
 DT 15-MAR-1991 (first entry)  
 XX Hypotensive oligopeptide.  
 XX Oral hypotensive; fig.  
 XX Ficus carica.  
 XX JP02282394-A.  
 PN 19-NOV-1990.  
 XX 24-APR-1989; 89JP-00104265.  
 XX 24-APR-1989; 89JP-00104265.  
 XX (AGEN ) AGENCY OF IND SCI & TECHNOLOGY.  
 PA (SHOS ) SHOWA SANGYO CO.  
 XX WPI; 1991-004480/01.  
 XX New oligopeptide hypotensive drug - based on alanine, valine, asparagine,  
 PT proline, isoleucine and arginine.  
 XX Claim 1; Page 843; 9pp; Japanese.  
 XX Peptides may be derived from extract of fig, and are hypotensive agents.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 XX Sequence 3 AA;  
 SQ  
 Query Match 55.6%; Score 5; DB 2; Length 3;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+06;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 R 1  
 Db 3 R 3  
 RESULT 25  
 AAR32271  
 ID AAR32271 standard; peptide; 3 AA.  
 AC AAR32271;  
 XX 01-APR-1993 (first entry)  
 DE Soybean glycinin derived hypotensive #13.  
 XX soybean; hypotensive; hypertension; high blood pressure.  
 KW Glycine max.  
 OS JP04297493-A.  
 PN 21-OCT-1992.  
 XX 13-FEB-1991; 91JP-00105245.  
 XX 13-FEB-1991; 91JP-00105245.  
 XX (AJIN ) AJINOMOTO KK.  
 PA WPI; 1992-401807/49.  
 XX Novel peptide(s) derived from soybean glycinin - are useful as  
 PT hypotensive agents.  
 XX Claim 1; Page 1; 5pp; Japanese.  
 PS This peptide is one of 17 derived from soybean glycinin. It is useful as  
 CC

CC a hypotensive agent in pharmaceuticals or food. It can be administered  
 CC orally or parenterally at a daily dose of 0.001mg to 5g. It may be  
 CC prepared by solid phase synthesis  
 XX Sequence 3 AA;  
 SQ  
 Query Match 55.6%; Score 5; DB 2; Length 3;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+06;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 R 1  
 Db 2 R 2  
 RESULT 26  
 AAR36707  
 ID AAR36707 standard; peptide; 3 AA.  
 XX AC AAR36707;  
 XX 25-MAR-2003 (revised)  
 DT 26-AUG-1993 (first entry)  
 XX Adhesion formation prevention RGD-contg. peptide.  
 DE Tissue repair; peritoneum; surgery; post-surgically; inhibition;  
 KW platelet aggregation; cardiovascular; orthopedic; thoracic; ophthalmic;  
 KW CNS; use.  
 XX Synthetic.  
 OS WO9308818-A1.  
 PN 13-MAY-1993.  
 PD 06-NOV-1992; 92WO-US009494.  
 XX 07-NOV-1991; 91US-00789231.  
 XX (UYSC-) UNIV SOUTHERN CALIFORNIA.  
 XX Dizerega GS, Rodgers KE;  
 XX WPI; 1993-167381/20.  
 XX Prevention of adhesion formation, partic. post-surgically - comprises  
 PT administering a RGD-contg. peptide for a time sufficient to permit tissue  
 PT repair.  
 XX Example; Page 18; 22pp; English.  
 PS The sequence is that of an RGD-contg. peptide which is used in a method  
 CC for prevention of adhesion formation for a time sufficient to permit  
 CC tissue repair. The method is used for minimising or preventing adhesion  
 CC formation, partic. in the peritoneum following surgery, but also for e.g.  
 CC cardiovascular, orthopedic, thoracic, ophthalmic, CNS and other uses. In  
 CC addn., the peptide inhibits platelet aggregation and does not induce  
 CC inflammation or trauma at the site of administration. (Updated on 25-MAR-  
 CC 2003 to correct PN field.)  
 XX  
 SQ Sequence 3 AA;  
 Query Match 55.6%; Score 5; DB 2; Length 3;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+06;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 R 1  
 Db 1 R 1  
 RESULT 27

OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers
FT	Misc-difference 3
FT	/note= "Leu185Pro"
XX	
PN	WO9312134-Al.
XX	
PD	24-JUN-1993.
XX	
PF	08-DEC-1992; 92WO-US010536.
XX	
PR	11-DEC-1991; 91US-00805123.
XX	(HARD ) HARVARD COLLEGE.
PA	Dryja TP, Berson EL;
PI	WPI; 1993-214088/26.
XX	N-PSDB; AAQ43556.
DR	
XX	Probe or primer contg. sequence of human retinal degeneration slow protein mutant - used to diagnose hereditary retinal degenerative diseases.
PT	
PT	Disclosure; Fig 13C; 56pp; English.
XX	
PS	The sequences given in AAR38485-87 represent regions of the human retinal degeneration slow (RDS) protein which contain mutations. These mutations of the human RDS protein cosegregate with autosomal dominant retinitis pigmentosa. The RDS gene sequence was isolated via the murine rds gene and has been mapped to chromosome 6p. The murine rds gene is a semidominant mutation with a phenotype of abnormal development of rod and cone photoreceptors, followed by their slow degeneration. The DNA encoding the mutation containing regions was isolated by polymerase chain reaction (PCR) using primers derived from RDS gene intron flanking sequences. (Updated on 25-MAR-2003 to correct FN field.)
CC	
XX	Sequence 3 AA;
SQ	
	Query Match 55.6%; Score 5; DB 2; Length 3;
	Best Local Similarity 100.0%; Pred. No. 1.9e+06; Indels 0; Gaps 0;
	Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 R 1
Db	1 R 1
RESULT 29	
AAR42569	
ID	AAR42569 standard; peptide; 3 AA.
XX	
AC	AAR42569;
XX	
DT	25-MAR-2003 (revised)
DT	22-JUN-1994 (first entry)
XX	
DE	Peptide corresponding to pseudo-substrate region of zeta-PKC.
XX	
KW	Zeta-protein kinase C inhibitor; zeta-PKC; pseudosubstrate; tumour;
KW	hyperproliferative disorders; psoriasis; viral infection; HIV.
XX	
OS	Synthetic.
XX	
FH	Key Location/Qualifiers
FT	Modified-site 1
FT	/note= "can be N-acetylated"
XX	
PN	WO9320101-Al.
XX	
XX	14-OCT-1993.
XX	



PF 02-APR-1993; 93WO-EP000816.  
 PR 06-APR-1992; 92EP-00500034.  
 XX (GLAX ) GLAXO SA.  
 XX Diaz-Meco Conde MT, Moscat Guillen J;  
 XX WPI; 1993-336831/42.  
 XX Peptide(s) corresp. to the pseudo-substrate region of zeta-PKC - used for  
 PT treatment of tumours, hyper-proliferative disorders and viral infections.  
 XX Claims 4 + 5; Page 43; 57pp; English.  
 XX The main claim refers to new peptides of formula X-Ala-Arg-Arg-J in which  
 CC X is H or one or more amino acids and J is OH or one or more amino acids,  
 CC the peptides containing a total of 3 to 15 amino acids. The present  
 CC peptide is a specifically claimed example of these new peptides. The  
 CC peptides are specific inhibitors of protein kinase C isotype zeta, i.e.  
 CC any subspecies of PKC which contains the specific autoinhibitory  
 CC pseudosubstrate domain RRGARRWK (Acc. No. AAR42573). This domain has  
 CC been found to be perfectly conserved in zeta-PKC variants isolated from a  
 CC number of different sources, including rat brain. The peptides are  
 CC usefully therapeutically for treating conditions where the underlying  
 CC aetiology is associated with zeta-PKC, including tumours,  
 CC hyperproliferative disorders (e.g. psoriasis) and viral infections (e.g.  
 CC HIV). (Updated on 25-MAR-2003 to correct PN field.)  
 XX Sequence 3 AA;  
 SQ Query Match 55.6%; Score 5; DB 2; Length 3;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+06;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 R 1  
 Db 2 R 2  
 RESULT 30  
 AAR30753  
 ID AAR30753 standard; peptide; 3 AA.  
 XX AAR30753;  
 XX 26-MAY-1993 (first entry)  
 XX IgG-mast cell reaction inhibitory peptide prepn. peptide.  
 XX Physiologically active; basophil; tripeptide.  
 XX Synthetic.  
 XX Key Location/Qualifiers  
 FT Modified-site 1  
 FT Modified-site 5 /note= "N-t-butoxycarbonyl-beta-benzyl-L-Asp"  
 FT Modified-site 5 /note= "N(G)-nitro-L-Arg benzyl ester"  
 XX JP04360899-A.  
 XX 14-DEC-1992.  
 XX 04-JUN-1991; 91JP-00159492.  
 XX 04-JUN-1991; 91JP-00159492.  
 XX (TANA ) TANABE SEIYAKU CO.  
 XX WPI; 1993-032711/04.  
 XX Prepn. of physiologically active penta:peptide - comprises fragment-

PT condensing specified carboxyl protected tri:peptide and amino-protected  
 XX di:peptide and removing protective gps.  
 XX Claim; Page 2; 8pp; Japanese.  
 XX The peptide is used in the prepn. (claimed) of a physiologically active  
 CC and prevents the IgE - mast cell (basophil) reaction  
 XX Sequence 3 AA;  
 SQ Query Match 55.6%; Score 5; DB 2; Length 3;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+06;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 R 1  
 Db 3 R 3  
 RESULT 31  
 AAR48960  
 ID AAR48960 standard; protein; 3 AA.  
 XX AAR48960;  
 XX 25-MAR-2003 (revised)  
 DT 12-SEP-1994 (first entry)  
 XX NL4-3 truncated GAG C-terminal peptide.  
 XX HIV-1; HXB2; antisense; sequence inversion; antisense virus; infection;  
 KW naturally occurring virus; NOV; translation; replication; infectivity;  
 KW hepatitis B; HIV-2; SIV; flip-over PCR.  
 XX Synthetic.  
 OS WO9403596-A1.  
 XX 17-FEB-1994.  
 PD 30-JUL-1993; 93WO-US007179.  
 PF 30-JUL-1992; 92US-00921104.  
 XX (UYHA-) UNIV HAWAII.  
 XX Hu W, Wang J;  
 XX WPI; 1994-065685/08.  
 DR N-PSDB; AAQ57688.  
 XX New antisense viruses and anti:sense-ribozyme viruses - used for treating  
 PT or preventing viral infections, partic. HIV-1, HIV-2 or SIV infection.  
 XX Disclosure; Page 108; 167pp; English.  
 XX This sequence is encoded by a PCR fragment of NL4-3 and represents the C-  
 CC terminal peptide fragment of the truncated GAG protein. The DNA encoding  
 CC this fragment was ligated into ClaI/SalI digested pX and the  
 CC corresponding plasmid was used to produce the antisense virus of the  
 CC invention. Antisense or truncated RNAs expressed by these viruses bind to  
 CC the mRNAs expressed by the naturally occurring viruses (NOVs) and prevent  
 CC the mRNAs from being translated into proteins, thereby preventing the NOV  
 CC from replicating. The antisense viruses maintain the infectivity of the  
 CC NOVs, allowing antisense RNAs to reach the mRNAs of the natural viruses.  
 CC Antisense viruses such as these may be used for treating or preventing a  
 CC viral infection, particularly HIV-1, HIV-2 or SIV infection or hepatitis  
 CC B infection. (Updated on 25-MAR-2003 to correct PN field.)  
 XX Sequence 3 AA;  
 SQ Query Match 55.6%; Score 5; DB 2; Length 3;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+06;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1  
Db 3 R 3

RESULT 32  
AAR46824  
ID AAR46824 standard; protein; 3 AA.

XX AC AAR46824;  
XX DT 25-MAR-2003 (revised)  
XX DT 19-AUG-1994 (first entry)  
XX DE Phytase derived peptide C-terminal (C phy).  
XX KW pH 2.5; acid phosphatase; Trichoderma; host; Aspergillus; phytic acid;  
XX KW phytate degrading enzyme; PDE; removal; inositol hexaphosphoric acid;  
XX KW plant; feed composition; filtration.  
XX OS Synthetic.  
XX PN WO9403612-A1.  
XX PD 17-FEB-1994.  
XX PF 30-JUL-1993; 93WO-FI000310.  
XX PR 31-JUL-1992; 92US-00923724.  
XX PA (ALKO-) ALKO LTD.  
XX PI Nevalainen HK, Paloheimo MT, Miettinen-Oinonen ASK, Torkkeli TK;  
XX PI Cantrell M, Piddington C, Rambosek JA, Turunen MK, Fagerstroem RB;  
XX WPI; 1994-065700/08.  
XX Compsns. contg. phytate degrading enzymes - obtd. by expression of their  
XX genes in Trichoderma, used partic. for producing animal feed compsns.  
XX Example 4; Page 45; 142pp; English.

XX The sequences given in AAR46793-824 are peptides derived from the phytase  
XX protein. The phytase protein may be used in the composition of the  
XX invention. The DNA encoding the phytase protein may be introduced into a  
XX Trichoderma host which then expresses it and the protein is collected  
XX from the culture medium. By using Trichoderma as a host for Aspergillus  
XX phytate degrading enzymes such as this, a totally different enzyme  
XX composition compared to that secreted from Aspergillus results. The  
XX enzyme composition can be used for removal of phytic acid or inositol  
XX hexaphosphoric acid from raw material, particularly plant material. The  
XX composition is used in feed compositions for animals. By using  
XX Trichoderma as a source of a composition containing phytate degrading  
XX enzymes some difficult downstream processing problems, eg. filtration,  
XX that occur with similar Aspergillus compositions are avoided and yields  
XX are improved. (Updated on 25-MAR-2003 to correct PN field.)

SQ Sequence 3 AA;  
Query Match 55.6%; Score 5; DB 2; Length 3;  
Best Local Similarity 100.0%; Pred. No. 1.9e+06;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1  
Db 1 R 1

RESULT 33  
AAR44666  
ID AAR44666 standard; peptide; 3 AA.

XX AAR44666;  
XX AC 20-JAN-1995 (first entry)  
XX DT Platelet aggregation or adhesion inhibitor - peptide 5.  
XX DE Platelet aggregation; adhesion; inhibitor; guest; host;  
XX KW beta-cyclodextrin; protease; resistance; degradation.  
XX OS Synthetic.  
XX PN JP06116289-A.  
XX PD 26-APR-1994.  
XX PF 09-OCT-1992; 92JP-00271294.  
XX PR 09-OCT-1992; 92JP-00271294.  
XX PA (FUJF) FUJI PHOTO FILM CO LTD.  
XX DR WPI; 1994-173759/21.  
XX Complex of adhesion peptide in host molecule e.g. beta-cyclodextrin -  
XX useful as platelet aggregation inhibitor which is resistant to protease  
XX degradation in-vivo.  
XX PS Disclosure; Page 4; 5pp; Japanese.  
XX CC A peptide complex contg. peptide 5 as guest mol. is useful as a platelet  
XX aggregation or adhesion inhibitor. It is hardly hydrolysed by protease in  
XX vivo and thus maintains its effect for a long period  
XX SQ Sequence 3 AA;  
Query Match 55.6%; Score 5; DB 2; Length 3;  
Best Local Similarity 100.0%; Pred. No. 1.9e+06;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1  
Db 1 R 1

RESULT 34  
AAR63264  
ID AAR63264 standard; peptide; 3 AA.  
XX AC AAR63264;  
XX DT 25-MAR-2003 (revised)  
XX DT 21-JUL-1995 (first entry)  
XX DE Thrombin inhibitor peptide, CSAP.  
XX KW Inhibitor; thrombin; chimeric molecule; fibrin-binding; antibody;  
XX KW hirudin; monoclonal; 59D8; fibrinopeptide B; coagulation;  
XX KW coronary stent implantation; adjunctive therapy; fibrinogen; haemorrhage.  
XX OS Synthetic.  
XX Key Location/Qualifiers  
XX FT Misc-difference 1 /note= "D form residue"  
XX FT Modified-site 3 /note= "opt. modified with -H to give ALD, or CH2Cl to  
XX FT give CMK"  
XX PN WO9425491-A1.  
XX PD 10-NOV-1994.  
XX

```

PF 03-MAY-1994; 94WO-US004881.
XX
PR 03-MAY-1993; 93US-00058699.
XX
PA (HARD ) HARVARD COLLEGE.
PA (UYEM-) UNIV EMORY.
XX
PI Haber E, Bode C, Runge M;
XX
DR WPI; 1994-358195/44.
XX
PT Fibrin-binding antibody linked to thrombin inhibitor - useful for
PT preventing blood coagulation by specifically targeting inhibitor to site
XX of thrombin activity.
PS Claim 7; Page 38; 53pp; English.
XX
CC This sequence represents an inhibitor of thrombin which was used in the
CC chimeric molecule of the invention. The chimeric molecule further
CC comprises a fibrin-binding antibody linked to the thrombin inhibitor
CC through a covalent linkage. The chimeric molecule allows fibrin-specific
CC antibody targeting of hirudin and other thrombin inhibitors, which is
CC more potent than thrombin on its own. The fibrin-specific antibody is
CC pref. the monoclonal antibody, 59D8. The epitope to which 59D8 binds
CC becomes available only after thrombin cleaves fibrinopeptide B. The
CC chimeric protein may be used for preventing coagulation of the blood.
CC Anti-thrombin targeting can be esp. useful in highly thrombogenic
CC situations such as coronary stent implantation and can be used as an
CC adjunctive therapy with highly selective thrombolytic agents. The
CC thrombin inhibitor is localised to sites of thrombin activity by the
CC antibody which binds to thrombin but does not cross react with uncleaved
CC fibrinogen. The selectivity of inhibition allows the total amount of
CC thrombin inhibitor used to be substantially reduced, resulting in a
CC reduced potential for generalised haemorrhaging. (Updated on 25-MAR-2003
XX to correct PN field.)
SQ Sequence 3 AA;

Query Match 55.6%; Score 5; DB 2; Length 3;
Best Local Similarity 100.0%; Pred. No. 1.9e+06;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1
Db 3 R 3

RESULT 35
AAR63263
ID AAR63263 standard; peptide; 3 AA.
XX
AC AAR63263;
XX
DT 25-MAR-2003 (revised)
DT 21-JUL-1995 (first entry)
XX
DE Thrombin inhibitor peptide #4.
XX
KW Inhibitor; thrombin; chimeric molecule; fibrin-binding; antibody;
KW hirudin; monoclonal; 59D8; fibrinopeptide B; coagulation;
KW coronary stent implantation; adjunctive therapy; fibrinogen; haemorrhage.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /label= H- (D) -Phe
FT Modified-site 3 /note= "boroArg-C10H16"
FT Modified-site 3 /note= "boroArg-C10H16"
XX
PN W09425491-A1.
XX
PD 10-NOV-1994.

XX 03-MAY-1994; 94WO-US004881.
XX
XX 03-MAY-1993; 93US-00058699.
XX
XX (HARD ) HARVARD COLLEGE.
XX (UYEM-) UNIV EMORY.
XX
XX Haber E, Bode C, Runge M;
XX
XX WPI; 1994-358195/44.
XX
XX Fibrin-binding antibody linked to thrombin inhibitor - useful for
XX preventing blood coagulation by specifically targeting inhibitor to site
XX of thrombin activity.
XX
XX Claim 7; Page 38; 53pp; English.
XX
XX This sequence represents an inhibitor of thrombin which was used in the
XX chimeric molecule of the invention. The chimeric molecule further
XX comprises a fibrin-binding antibody linked to the thrombin inhibitor
XX through a covalent linkage. The chimeric molecule allows fibrin-specific
XX antibody targeting of hirudin and other thrombin inhibitors, which is
XX more potent than thrombin on its own. The fibrin-specific antibody is
XX pref. the monoclonal antibody, 59D8. The epitope to which 59D8 binds
XX becomes available only after thrombin cleaves fibrinopeptide B. The
XX chimeric protein may be used for preventing coagulation of the blood.
XX Anti-thrombin targeting can be esp. useful in highly thrombogenic
XX situations such as coronary stent implantation and can be used as an
XX adjunctive therapy with highly selective thrombolytic agents. The
XX thrombin inhibitor is localised to sites of thrombin activity by the
XX antibody which binds to thrombin but does not cross react with uncleaved
XX fibrinogen. The selectivity of inhibition allows the total amount of
XX thrombin inhibitor used to be substantially reduced, resulting in a
XX reduced potential for generalised haemorrhaging. (Updated on 25-MAR-2003
XX to correct PN field.)
XX
XX Sequence 3 AA;

Query Match 55.6%; Score 5; DB 2; Length 3;
Best Local Similarity 100.0%; Pred. No. 1.9e+06;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1
Db 3 R 3

RESULT 36
AAR63261
ID AAR63261 standard; peptide; 3 AA.
XX
AC AAR63261;
XX
DT 25-MAR-2003 (revised)
DT 21-JUL-1995 (first entry)
XX
DE Thrombin inhibitor peptide #2.
XX
KW Inhibitor; thrombin; chimeric molecule; fibrin-binding; antibody;
KW hirudin; monoclonal; 59D8; fibrinopeptide B; coagulation;
KW coronary stent implantation; adjunctive therapy; fibrinogen; haemorrhage.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /label= Boc- (D) -Phe
FT Modified-site 3 /note= "boroArg-C10H16"
FT Modified-site 3 /note= "boroArg-C10H16"
XX
XX W09425491-A1.
XX
XX W09425491-A1.

```



XX PN WO9425491-A1.  
 XX XX 10-NOV-1994.  
 XX PD PA (EMBR-) EMBREX INC.  
 XX PF PA (GROP-) GROPEP PTY LTD.  
 XX PR PA (USDA ) US SEC OF AGRIC.  
 XX PI Francis GL, Walton PE, Ballard FJ, Mcmurtry JP, Phelps PV;  
 XX DR WPI; 1994-118144/14.  
 XX PT Increasing growth of birds - with insulin-like growth factor delivered to  
 XX PT the egg before hatching esp. for increasing wt. gain in chickens.  
 XX PS Claim 6; Page 37; 45pp; English.  
 XX CC Growth of birds is increased by (a) admin. to the bird, in ovo, insulin-  
 XX CC like growth factor (IGF)-1 or -2 or their active analogues; (b)  
 XX CC incubating to hatch and (c) growing the birds for at least 3 weeks after  
 XX CC hatch. IGF-1 analogues used pref. (i) lack 1-5 N-terminal amino acids or  
 XX CC (ii) have Glu(3) absent or replaced by Gly, Gln, Leu, Arg or Lys, partic.  
 XX CC Gly and the Thr normally adjacent to Glu can be replaced by Arg or Gly.  
 XX CC The IGF-1 analogue pref. comprises the N-terminal given in AAR51439-49.  
 XX CC IGF-2 analogues used pref. have Glu(2) or Glu(3) absent or replaced by  
 XX CC Gly, Gln, Leu, Arg or Lys, partic. Gly and the Thr normally adjacent to  
 XX CC Glu can be replaced by Arg or Gly. The IGF-2 analogue pref. comprises the  
 XX CC N-terminal given in AAR51450-53. Esp. Long R3 IGF-1, given in AAR51454 is  
 XX CC used. It is the full human IGF-1 sequence with Arg replacing Glu(3) and a  
 XX CC 13 amino acid N-terminal extension. (Updated on 25-MAR-2003 to correct PN  
 XX CC field.)  
 XX SQ Sequence 3 AA;  
 Query Match 55.6%; Score 5; DB 2; Length 3;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+06;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 R 1  
 Db 1 R 1  
 RESULT 40  
 AAR61094  
 ID AAR61094 standard; peptide; 3 AA.  
 XX AC AAR61094;  
 XX DT 21-MAY-1995 (first entry)  
 XX DE ACE-inhibiting tripeptide.  
 XX KW ACE; angiotensin converting enzyme inhibitor; fish meat alpha-1000;  
 XX KW hypotensive.  
 XX OS Synthetic.  
 XX PN JP06166697-A.  
 XX PD 14-JUN-1994.  
 XX PF 01-DEC-1992; 92JP-00343573.  
 XX PR 01-DEC-1992; 92JP-00343573.  
 XX PA (SENN-) SENMI EKESU KK.  
 XX DR WPI; 1994-230661/28.  
 XX PT New ACE-inhibiting di-, tri- and tetra-peptide(s) - obt'd. by treating  
 XX PT fish meat alpha-1000 peptide with ODS resin.  
 XX PS Claim 1; Page 2; 10pp; Japanese.

XX PN WO9425491-A1.  
 XX XX 10-NOV-1994.  
 XX PD PA (EMBR-) EMBREX INC.  
 XX PF PA (GROP-) GROPEP PTY LTD.  
 XX PR PA (USDA ) US SEC OF AGRIC.  
 XX PI Francis GL, Walton PE, Ballard FJ, Mcmurtry JP, Phelps PV;  
 XX DR WPI; 1994-118144/14.  
 XX PT Increasing growth of birds - with insulin-like growth factor delivered to  
 XX PT the egg before hatching esp. for increasing wt. gain in chickens.  
 XX PS Claim 6; Page 37; 45pp; English.  
 XX CC Growth of birds is increased by (a) admin. to the bird, in ovo, insulin-  
 XX CC like growth factor (IGF)-1 or -2 or their active analogues; (b)  
 XX CC incubating to hatch and (c) growing the birds for at least 3 weeks after  
 XX CC hatch. IGF-1 analogues used pref. (i) lack 1-5 N-terminal amino acids or  
 XX CC (ii) have Glu(3) absent or replaced by Gly, Gln, Leu, Arg or Lys, partic.  
 XX CC Gly and the Thr normally adjacent to Glu can be replaced by Arg or Gly.  
 XX CC The IGF-1 analogue pref. comprises the N-terminal given in AAR51439-49.  
 XX CC IGF-2 analogues used pref. have Glu(2) or Glu(3) absent or replaced by  
 XX CC Gly, Gln, Leu, Arg or Lys, partic. Gly and the Thr normally adjacent to  
 XX CC Glu can be replaced by Arg or Gly. The IGF-2 analogue pref. comprises the  
 XX CC N-terminal given in AAR51450-53. Esp. Long R3 IGF-1, given in AAR51454 is  
 XX CC used. It is the full human IGF-1 sequence with Arg replacing Glu(3) and a  
 XX CC 13 amino acid N-terminal extension. (Updated on 25-MAR-2003 to correct PN  
 XX CC field.)  
 XX SQ Sequence 3 AA;  
 Query Match 55.6%; Score 5; DB 2; Length 3;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+06;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 R 1  
 Db 3 R 3  
 RESULT 39  
 AAR51440  
 ID AAR51440 standard; peptide; 3 AA.  
 XX AC AAR51440;  
 XX DT 25-MAR-2003 (revised)  
 XX DT 27-OCT-1994 (first entry)  
 XX DE IGF-1 analogue N-terminal.  
 XX KW Insulin-like growth factor; IGF-1; IGF-2; bird; chicken; egg; in ovo;  
 XX KW growth; promotion; increase; long R3 IGF-1; LR3 IGF-1.  
 XX OS Homo sapiens.  
 XX PN WO9406445-A1.  
 XX PD 31-MAR-1994.  
 XX PF 02-SEP-1993; 93WO-US008279.

Fibrin-binding antibody linked to thrombin inhibitor - useful for  
 preventing blood coagulation by specifically targeting inhibitor to site  
 of thrombin activity.  
 Claim 7; Page 38; 53pp; English.  
 This sequence represents an inhibitor of thrombin which was used in the  
 chimeric molecule of the invention. The chimeric molecule further  
 comprises a fibrin-binding antibody linked to the thrombin inhibitor  
 through a covalent linkage. The chimeric molecule allows fibrin-specific  
 antibody targeting of hirudin and other thrombin inhibitors, which is  
 more potent than thrombin on its own. The fibrin-specific antibody is  
 pref. the monoclonal antibody, 59D8. The epitope to which 59D8 binds  
 becomes available only after thrombin cleaves fibrinopeptide B. The  
 chimeric protein may be used for preventing coagulation of the blood.  
 Anti-thrombin targeting can be esp. useful in highly thrombogenic  
 situations such as coronary stent implantation and can be used as an  
 adjunctive therapy with highly selective thrombolytic agents. The  
 thrombin inhibitor is localised to sites of thrombin activity by the  
 antibody which binds to thrombin but does not cross react with uncleaved  
 fibrinogen. The selectivity of inhibition allows a total amount of  
 thrombin inhibitor used to be substantially reduced, resulting in a  
 reduced potential for generalised haemorrhaging. (Updated on 25-MAR-2003  
 to correct PN field.)

```

XX CC A total of ten di-, tri- or tetrapeptides are claimed which have
CC CC angiotensin converting enzyme inhibiting activity and which are useful as
CC CC hypotensives. The present sequence is one of the ten. The peptides are
CC CC obtained by treating fish meat alpha-1000 peptide with ODS resin. This
CC CC peptide has an ACE inhibiting IC50 value of 205.6 micromolar
XX SQ Sequence 3 AA;

Query Match 55.6%; Score 5; DB 2; Length 3;
Best Local Similarity 100.0%; Pred. No. 1.9e+06;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1
Db 1 R 1

RESULT 41
AAR61092
ID AAR61092 standard; peptide; 3 AA.
XX AC AAR61092;
XX DT 21-MAY-1995 (first entry)
XX DE ACE-inhibiting tripeptide.
XX KW ACE; angiotensin converting enzyme inhibitor; fish meat alpha-1000;
XX KW hypotensive.
XX OS Synthetic.
XX PN JP06166697-A.
XX PD 14-JUN-1994.
XX PF 01-DEC-1992; 92JP-00343573.
XX PR 01-DEC-1992; 92JP-00343573.
XX PA (SENM-) SENMI EKESU KK.
XX DR WPI; 1994-230661/28.
XX DT New ACE-inhibiting di-, tri- and tetra:peptide(s) - obtd. by treating
XX PT fish meat alpha-1000 peptide with ODS resin.
XX PS Claim 1; Page 2; 10pp; Japanese.
XX CC A total of ten di-, tri- or tetrapeptides are claimed which have
XX CC angiotensin converting enzyme inhibiting activity and which are useful as
XX CC hypotensives. The present sequence is one of the ten. The peptides are
XX CC obtained by treating fish meat alpha-1000 peptide with ODS resin. This
XX CC peptide has an ACE inhibiting IC50 value of 20.0 micromolar
XX SQ Sequence 3 AA;

Query Match 55.6%; Score 5; DB 2; Length 3;
Best Local Similarity 100.0%; Pred. No. 1.9e+06;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1
Db 2 R 2

RESULT 42
AAR61093
ID AAR61093 standard; peptide; 3 AA.
XX AC AAR61093;
XX DT 21-MAY-1995 (first entry)
XX DE ACE-inhibiting tripeptide.
XX KW ACE; angiotensin converting enzyme inhibitor; fish meat alpha-1000;
XX KW hypotensive.
XX OS Synthetic.
XX PN JP06166697-A.
XX PD 14-JUN-1994.
XX PF 01-DEC-1992; 92JP-00343573.
XX PR 01-DEC-1992; 92JP-00343573.
XX PA (SENM-) SENMI EKESU KK.
XX DR WPI; 1994-230661/28.
XX DT New ACE-inhibiting di-, tri- and tetra:peptide(s) - obtd. by treating
XX PT fish meat alpha-1000 peptide with ODS resin.
XX PS Claim 1; Page 2; 10pp; Japanese.
XX CC A total of ten di-, tri- or tetrapeptides are claimed which have
XX CC angiotensin converting enzyme inhibiting activity and which are useful as
XX CC hypotensives. The present sequence is one of the ten. The peptides are
XX CC obtained by treating fish meat alpha-1000 peptide with ODS resin. This
XX CC peptide has an ACE inhibiting IC50 value of 20.0 micromolar
XX SQ Sequence 3 AA;

Query Match 55.6%; Score 5; DB 2; Length 3;
Best Local Similarity 100.0%; Pred. No. 1.9e+06;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1
Db 2 R 2

RESULT 43
AAR73934
ID AAR73934 standard; peptide; 3 AA.
XX AC AAR73934;
XX DT 14-JUL-1995 (first entry)
XX DE Novel tripeptide inhibitor of angiotensin transferase.
XX KW angiotensin transferase; inhibitor; hypertension; casein; treatment;
XX KW enzymatic hydrolysis.
XX OS Synthetic.
XX PN JP06277091-A.
XX PD 04-OCT-1994.
XX PF 26-MAR-1993; 93JP-00092553.
XX PR 26-MAR-1993; 93JP-00092553.
XX PA (NISY ) NIPPON SYNTHETIC CHEM IND CO.

```

```

DT 21-MAY-1995 (first entry)
XX ACB-inhibiting tripeptide.
XX DE ACE-inhibiting tripeptide.
XX KW ACE; angiotensin converting enzyme inhibitor; fish meat alpha-1000;
XX KW hypotensive.
XX OS Synthetic.
XX PN JP06166697-A.
XX PD 14-JUN-1994.
XX PF 01-DEC-1992; 92JP-00343573.
XX PR 01-DEC-1992; 92JP-00343573.
XX PA (SENM-) SENMI EKESU KK.
XX DR WPI; 1994-230661/28.
XX DT New ACE-inhibiting di-, tri- and tetra:peptide(s) - obtd. by treating
XX PT fish meat alpha-1000 peptide with ODS resin.
XX PS Claim 1; Page 2; 10pp; Japanese.
XX CC A total of ten di-, tri- or tetrapeptides are claimed which have
XX CC angiotensin converting enzyme inhibiting activity and which are useful as
XX CC hypotensives. The present sequence is one of the ten. The peptides are
XX CC obtained by treating fish meat alpha-1000 peptide with ODS resin. This
XX CC peptide has an ACE inhibiting IC50 value of 330 micromolar
XX SQ Sequence 3 AA;

Query Match 55.6%; Score 5; DB 2; Length 3;
Best Local Similarity 100.0%; Pred. No. 1.9e+06;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1
Db 1 R 1

RESULT 43
AAR73934
ID AAR73934 standard; peptide; 3 AA.
XX AC AAR73934;
XX DT 14-JUL-1995 (first entry)
XX DE Novel tripeptide inhibitor of angiotensin transferase.
XX KW angiotensin transferase; inhibitor; hypertension; casein; treatment;
XX KW enzymatic hydrolysis.
XX OS Synthetic.
XX PN JP06277091-A.
XX PD 04-OCT-1994.
XX PF 26-MAR-1993; 93JP-00092553.
XX PR 26-MAR-1993; 93JP-00092553.
XX PA (NISY ) NIPPON SYNTHETIC CHEM IND CO.

```

XX WPI; 1994-353767/44.  
 XX DR Novel tri-peptide inhibitor of angiotensin converting enzyme - obtained  
 XX PT by enzymatic hydrolysis of protein, e.g. casein.  
 XX PS Claim 1; Page 2; 4pp; Japanese.  
 XX SS AAR73934 is a novel tripeptide useful for the treatment of hypertension.  
 CC It inhibits angiotensin transferase and is obtd. by the enzymatic  
 CC hydrolysis of proteins using thermolysin, in partic. casein. It is also  
 CC useful in the treatment of angina pectoris, ischaemic cardiac  
 CC insufficiency, and myocardial infarction  
 XX SQ Sequence 3 AA;  
 Query Match 55.6%; Score 5; DB 2; Length 3;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+06;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 R 1  
 Db 2 R 2  
 RESULT 44  
 AAR57458  
 ID AAR57458 standard; protein; 3 AA.  
 XX AC AAR57458;  
 XX DT 28-FEB-1995 (first entry)  
 XX DE Lactoferrin derived peptide #21.  
 XX KW Lactoferrin; chemical; enzymatic; hydrolysis; antimicrobial; antiseptic;  
 XX KW ischaemic disease.  
 XX OS Mus musculus.  
 XX PN JP06172200-A.  
 XX PD 21-JUN-1994.  
 XX PF 08-DEC-1992; 92JP-00327738.  
 XX PR 08-DEC-1992; 92JP-00327738.  
 XX PA (MORG ) MORINAGA MILK IND CO LTD.  
 XX DR WPI; 1994-238662/29.  
 XX PT Brain protectant for preventing ischaemic diseases without side effects -  
 XX PT comprising 31 specified peptide(s), prepd. by lactoferrin hydrolysis.  
 XX PS Disclosure; Page 9; 11pp; Japanese.  
 CC The sequences given in AAR57438-68 represent fragments of lactoferrin  
 CC which were derived from the full length protein by chemical or enzyme  
 CC hydrolysis. These peptides have brain protecting properties, as well as  
 CC anti-microbial activity. Compositions containing these peptides may be  
 CC prepared with out the addition of antiseptics, and may be administered at  
 CC doses of at least 10 mg for parenteral administration and 100 mg for oral  
 CC administration. These peptides are stable, heat resistant, water soluble  
 CC and may be used for the prevention of ischaemic diseases without side  
 CC effects  
 XX SQ Sequence 3 AA;  
 Query Match 55.6%; Score 5; DB 2; Length 3;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+06;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 R 1  
 Db 2 R 2  
 RESULT 45  
 AAR57450  
 ID AAR57450 standard; protein; 3 AA.  
 XX AC AAR57450;  
 XX DT 28-FEB-1995 (first entry)  
 XX DE Lactoferrin derived peptide #13.  
 XX KW Lactoferrin; chemical; enzymatic; hydrolysis; antimicrobial; antiseptic;  
 XX KW ischaemic disease.  
 XX OS Mus musculus.  
 XX PN JP06172200-A.  
 XX PD 21-JUN-1994.  
 XX PF 08-DEC-1992; 92JP-00327738.  
 XX PR 08-DEC-1992; 92JP-00327738.  
 XX PA (MORG ) MORINAGA MILK IND CO LTD.  
 XX DR WPI; 1994-238662/29.  
 XX PT Brain protectant for preventing ischaemic diseases without side effects -  
 XX PT comprising 31 specified peptide(s), prepd. by lactoferrin hydrolysis.  
 XX PS Disclosure; Page 8; 11pp; Japanese.  
 CC The sequences given in AAR57438-68 represent fragments of lactoferrin  
 CC which were derived from the full length protein by chemical or enzyme  
 CC hydrolysis. These peptides have brain protecting properties, as well as  
 CC anti-microbial activity. Compositions containing these peptides may be  
 CC prepared with out the addition of antiseptics, and may be administered at  
 CC doses of at least 10 mg for parenteral administration and 100 mg for oral  
 CC administration. These peptides are stable, heat resistant, water soluble  
 CC and may be used for the prevention of ischaemic diseases without side  
 CC effects  
 XX SQ Sequence 3 AA;  
 Query Match 55.6%; Score 5; DB 2; Length 3;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+06;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 R 1  
 Db 3 R 3  
 RESULT 46  
 AAR48523  
 ID AAR48523 standard; peptide; 3 AA.  
 XX AC AAR48523;  
 XX DT 25-MAR-2003 (revised)  
 XX DT 10-AUG-1994 (first entry)  
 XX DE Lactoferrin derived peptide #17.  
 XX KW Decomposition; lactoferrin; digestion; enzyme; pepsin; trypsin;  
 XX KW antioxidant; oxidation; inhibitor; vitamin E; ascorbic acid; vitamin A;  
 XX KW beta-carotene; superoxide dismutase; coenzyme Q; lipid oxidation;  
 XX KW foodstuff; drugs; health food; toiletries; cosmetics.

Qy 1 R 1  
 Db 1 R 1  
 RESULT 45  
 AAR57450  
 ID AAR57450 standard; protein; 3 AA.  
 XX AC AAR57450;  
 XX DT 28-FEB-1995 (first entry)  
 XX DE Lactoferrin derived peptide #13.  
 XX KW Lactoferrin; chemical; enzymatic; hydrolysis; antimicrobial; antiseptic;  
 XX KW ischaemic disease.  
 XX OS Mus musculus.  
 XX PN JP06172200-A.  
 XX PD 21-JUN-1994.  
 XX PF 08-DEC-1992; 92JP-00327738.  
 XX PR 08-DEC-1992; 92JP-00327738.  
 XX PA (MORG ) MORINAGA MILK IND CO LTD.  
 XX DR WPI; 1994-238662/29.  
 XX PT Brain protectant for preventing ischaemic diseases without side effects -  
 XX PT comprising 31 specified peptide(s), prepd. by lactoferrin hydrolysis.  
 XX PS Disclosure; Page 8; 11pp; Japanese.  
 CC The sequences given in AAR57438-68 represent fragments of lactoferrin  
 CC which were derived from the full length protein by chemical or enzyme  
 CC hydrolysis. These peptides have brain protecting properties, as well as  
 CC anti-microbial activity. Compositions containing these peptides may be  
 CC prepared with out the addition of antiseptics, and may be administered at  
 CC doses of at least 10 mg for parenteral administration and 100 mg for oral  
 CC administration. These peptides are stable, heat resistant, water soluble  
 CC and may be used for the prevention of ischaemic diseases without side  
 CC effects  
 XX SQ Sequence 3 AA;  
 Query Match 55.6%; Score 5; DB 2; Length 3;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+06;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 R 1  
 Db 3 R 3  
 RESULT 46  
 AAR48523  
 ID AAR48523 standard; peptide; 3 AA.  
 XX AC AAR48523;  
 XX DT 25-MAR-2003 (revised)  
 XX DT 10-AUG-1994 (first entry)  
 XX DE Lactoferrin derived peptide #17.  
 XX KW Decomposition; lactoferrin; digestion; enzyme; pepsin; trypsin;  
 XX KW antioxidant; oxidation; inhibitor; vitamin E; ascorbic acid; vitamin A;  
 XX KW beta-carotene; superoxide dismutase; coenzyme Q; lipid oxidation;  
 XX KW foodstuff; drugs; health food; toiletries; cosmetics.

XX OS Bos taurus.  
 XX WO9403555-A1.  
 XX 17-FEB-1994.  
 XX 04-AUG-1993; 93WO-JP001090.  
 XX 07-AUG-1992; 92JP-00211335.  
 XX (MORG ) MORINAGA MILK IND CO LTD.  
 XX Tomita M, Shimamura S, Kawase K, Fukuwatari Y, Takase M;  
 XX Bellamy WR, Yamauchi K, Wakabayashi H, Tokida Y;  
 XX WPI; 1994-065650/08.  
 XX Antioxidant peptide lactoferrin decomposition product - prevents  
 PT oxidation of lipid(s) in foodstuffs and drugs without affecting their  
 PT taste.  
 XX Claim 3; Page 32; 47pp; Japanese.  
 XX The sequences given in AAR48507-37 are peptides derived by the  
 CC decomposition of lactoferrin, pref. by digestion with an enzyme, eg.  
 CC pepsin or trypsin. These peptides may be used in an antioxidant  
 CC composition which may also contain an oxidation inhibitor such as vitamin  
 CC E, ascorbic acid, vitamin A, beta-carotene, superoxide dismutase or  
 CC coenzyme Q. The antioxidant prevents lipid oxidation in foodstuffs,  
 CC drugs, health foods, toiletries and cosmetics. (Updated on 25-MAR-2003 to  
 CC correct PN field.)  
 XX Sequence 3 AA;  
 SQ Query Match 55.6%; Score 5; DB 2; Length 3;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+06;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 R 1  
 Db 3 R 3  
 RESULT 47  
 AAR48527  
 ID AAR48527 standard; peptide; 3 AA.  
 AC AAR48527;  
 XX 25-MAR-2003 (revised)  
 DT 10-AUG-1994 (first entry)  
 XX Lactoferrin derived peptide #21.  
 DE Decomposition; lactoferrin; digestion; enzyme; pepsin; trypsin;  
 KW antioxidant; oxidation; inhibitor; vitamin E; ascorbic acid; vitamin A;  
 KW beta-carotene; superoxide dismutase; coenzyme Q; lipid oxidation;  
 KW foodstuff; drugs; health food; toiletries; cosmetics.  
 XX Bos taurus.  
 OS WO9403555-A1.  
 XX 17-FEB-1994.  
 XX 04-AUG-1993; 93WO-JP001090.  
 XX 07-AUG-1992; 92JP-00211335.  
 XX (MORG ) MORINAGA MILK IND CO LTD.  
 XX Tomita M, Shimamura S, Kawase K, Fukuwatari Y, Takase M;  
 XX WPI; 1994-065650/08.  
 XX Antioxidant peptide lactoferrin decomposition product - prevents  
 PT oxidation of lipid(s) in foodstuffs and drugs without affecting their  
 PT taste.  
 XX Claim 3; Page 30; 47pp; Japanese.  
 XX The sequences given in AAR48507-37 are peptides derived by the  
 CC decomposition of lactoferrin, pref. by digestion with an enzyme, eg.  
 CC pepsin or trypsin. These peptides may be used in an antioxidant  
 CC composition which may also contain an oxidation inhibitor such as vitamin  
 CC E, ascorbic acid, vitamin A, beta-carotene, superoxide dismutase or  
 CC coenzyme Q. The antioxidant prevents lipid oxidation in foodstuffs,  
 CC drugs, health foods, toiletries and cosmetics. (Updated on 25-MAR-2003 to  
 CC correct PN field.)  
 XX Sequence 3 AA;  
 SQ Query Match 55.6%; Score 5; DB 2; Length 3;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+06;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 R 1  
 Db 3 R 3

PI Bellamy WR, Yamauchi K, Wakabayashi H, Tokida Y;  
 XX WPI; 1994-065650/08.  
 XX Antioxidant peptide lactoferrin decomposition product - prevents  
 PT oxidation of lipid(s) in foodstuffs and drugs without affecting their  
 PT taste.  
 XX Claim 3; Page 34; 47pp; Japanese.  
 XX The sequences given in AAR48507-37 are peptides derived by the  
 CC decomposition of lactoferrin, pref. by digestion with an enzyme, eg.  
 CC pepsin or trypsin. These peptides may be used in an antioxidant  
 CC composition which may also contain an oxidation inhibitor such as vitamin  
 CC E, ascorbic acid, vitamin A, beta-carotene, superoxide dismutase or  
 CC coenzyme Q. The antioxidant prevents lipid oxidation in foodstuffs,  
 CC drugs, health foods, toiletries and cosmetics. (Updated on 25-MAR-2003 to  
 CC correct PN field.)  
 XX Sequence 3 AA;  
 SQ Query Match 55.6%; Score 5; DB 2; Length 3;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+06;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 R 1  
 Db 1 R 1  
 RESULT 48  
 AAR48519  
 ID AAR48519 standard; peptide; 3 AA.  
 XX AAR48519;  
 XX 25-MAR-2003 (revised)  
 DT 10-AUG-1994 (first entry)  
 XX Lactoferrin derived peptide #13.  
 DE Decomposition; lactoferrin; digestion; enzyme; pepsin; trypsin;  
 KW antioxidant; oxidation; inhibitor; vitamin E; ascorbic acid; vitamin A;  
 KW beta-carotene; superoxide dismutase; coenzyme Q; lipid oxidation;  
 KW foodstuff; drugs; health food; toiletries; cosmetics.  
 XX Bos taurus.  
 OS WO9403555-A1.  
 XX 17-FEB-1994.  
 XX 04-AUG-1993; 93WO-JP001090.  
 XX 07-AUG-1992; 92JP-00211335.  
 XX (MORG ) MORINAGA MILK IND CO LTD.  
 XX Tomita M, Shimamura S, Kawase K, Fukuwatari Y, Takase M;  
 XX Bellamy WR, Yamauchi K, Wakabayashi H, Tokida Y;  
 XX WPI; 1994-065650/08.  
 XX Antioxidant peptide lactoferrin decomposition product - prevents  
 PT oxidation of lipid(s) in foodstuffs and drugs without affecting their  
 PT taste.  
 XX Claim 3; Page 30; 47pp; Japanese.  
 XX The sequences given in AAR48507-37 are peptides derived by the  
 CC decomposition of lactoferrin, pref. by digestion with an enzyme, eg.  
 CC pepsin or trypsin. These peptides may be used in an antioxidant  
 CC composition which may also contain an oxidation inhibitor such as vitamin



CC E, ascorbic acid, vitamin A, beta-carotene, superoxidase dismutase or  
CC coenzyme Q. The antioxidant prevents lipid oxidation in foodstuffs,  
CC drugs, health foods, toiletries and cosmetics. (Updated on 25-MAR-2003 to  
CC correct PN field.)  
XX  
XX  
SQ Sequence 3 AA;

Query Match

Best Local Similarity 55.6%; Score 5; DB 2; Length 3;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 R 1

Db 3 R 3

RESULT 49

AAR58575

ID AAR58575 standard; peptide; 3 AA.

XX AC AAR58575;

XX DT 26-APR-1995 (first entry)

DE Angiotensin I converting enzyme inhibitory tripeptide FRP.

XX KW angiotensin converting enzyme; inhibitor; food ingredient.

OS Synthetic.

XX PN JP06220088-A.

XX PD 09-AUG-1994.

XX PF 22-JAN-1993; 93JP-00025977.

XX PR 22-JAN-1993; 93JP-00025977.

XX PA (ASAH ) ASahi KASEI KOGYO KK.

XX DR WPI; 1994-290911/36.

XX PT New tri:peptide(s) - inhibit angiotensin I converting enzyme.

XX PS Claim 1; Page 2; 4pp; Japanese.

CC This is one of thirteen claimed tripeptides (AAR58569-R58581) which

CC inhibit angiotensin I converting enzyme (with IC50 of 2.9-186.2

CC micromolar). The tripeptides are incorporated into food, e.g. hamburgers

XX SQ Sequence 3 AA;

Query Match

Best Local Similarity 55.6%; Score 5; DB 2; Length 3;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 R 1

Db 2 R 2

RESULT 50

AAR58571

ID AAR58571 standard; peptide; 3 AA.

XX AC AAR58571;

XX DT 26-APR-1995 (first entry)

DE Angiotensin I converting enzyme inhibitory tripeptide LRY.

XX KW angiotensin converting enzyme; inhibitor; food ingredient.

OS Synthetic.

XX PN JP06220088-A.

XX PD 09-AUG-1994.

XX PF 22-JAN-1993; 93JP-00025977.

XX PR 22-JAN-1993; 93JP-00025977.

XX PA (ASAH ) ASahi KASEI KOGYO KK.

XX DR WPI; 1994-290911/36.

XX PT New tri:peptide(s) - inhibit angiotensin I converting enzyme.

XX PS Claim 1; Page 2; 4pp; Japanese.

XX CC This is one of thirteen claimed tripeptides (AAR58569-R58581) which

XX CC inhibit angiotensin I converting enzyme (with IC50 of 2.9-186.2

XX CC micromolar). The tripeptides are incorporated into food, e.g. hamburgers

XX SQ Sequence 3 AA;

Query Match

Best Local Similarity 55.6%; Score 5; DB 2; Length 3;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 R 1

Db 2 R 2

Search completed: January 25, 2006, 18:38:39

Job time : 89.5 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:34:13 ; Search time 12.5 Seconds  
(without alignments)  
38.487 Million cell updates/sec

Title: US-10-771-242-293

Perfect score: 9

Sequence: 1 RXXX 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5	55.6	3	A23565	R-phycoerythrin al
2	5	55.6	3	PQ0010	angiotensin-conver
3	5	55.6	4	ECXAA	antho-RFamide-neu
4	5	55.6	4	A02147	phagocytosis-stimu
5	5	55.6	4	D41654	hypothetical prote
6	5	55.6	4	I40870	phospholipase C [E
7	5	55.6	4	A25844	antho-RF amide neu
8	5	55.6	4	I61883	protamine p1 4 ora
9	5	55.6	4	S47552	ubiquitin - rat
10	5	55.6	4	I37013	protamine p1 - Cer
11	5	55.6	4	I84439	protamine p1 - sav
12	5	55.6	4	A35779	neuropeptide Antho
13	5	55.6	4	A60418	FMRFamide - polych
14	5	55.6	4	PT0721	T-cell receptor be
15	5	55.6	4	ECNK	cardioexcitatory n
16	5	55.6	5	1	protocollin - Americ
17	5	55.6	5	JN0862	peptidyl-di-peptida
18	5	55.6	5	I40702	primase - Citrobac
19	5	55.6	5	A44955	alkanal monooxygen
20	5	55.6	5	I39964	ribosomal protein
21	5	55.6	5	I39966	ribosomal protein
22	5	55.6	5	I39965	major protein anti
23	5	55.6	5	D60274	R-phycoerythrin al
24	5	55.6	5	B22565	R-phycoerythrin ga
25	5	55.6	5	T22565	hypothetical prote
26	5	55.6	5	T14910	neuropeptide - sea
27	5	55.6	5	A60803	hypothetical prote
28	5	55.6	5	S53595	Ig heavy chain CRD
29	5	55.6	5	PT0295	

30	5	55.6	5	2	S62883	seminal plasma pro
31	5	55.6	5	2	PT0513	T-cell receptor be
32	5	55.6	5	2	PT0525	T-cell receptor be
33	5	55.6	5	2	PT0597	T-cell receptor be
34	5	55.6	5	2	PT0672	T-cell receptor be
35	5	55.6	5	2	PT0553	T-cell receptor be
36	5	55.6	5	2	PT0695	T-cell receptor be
37	5	55.6	5	2	PT0577	T-cell receptor be
38	5	55.6	5	2	PT0572	T-cell receptor be
39	5	55.6	5	2	PT0700	T-cell receptor be
40	5	55.6	5	2	A60411	protocollin - Atlant
41	5	55.6	5	2	PT0608	T-cell receptor be
42	5	55.6	5	2	PT0565	T-cell receptor be
43	5	55.6	6	2	A35890	RNA-directed DNA p
44	5	55.6	6	2	A37765	hypothetical prote
45	5	55.6	6	2	C22565	R-phycoerythrin be
46	5	55.6	6	2	PQ0008	angiotensin-conver
47	5	55.6	6	2	A60494	antineoplastic gly
48	5	55.6	6	2	I51434	H4 histone - Afric
49	5	55.6	6	2	I37027	protamine p1 - gor
50	5	55.6	6	2	A11490	pyruvate kinase (E
51	5	55.6	6	2	B56979	collagen alpha 1(I
52	5	55.6	6	2	A19780	transferrin - bovi
53	5	55.6	6	2	B26206	alpha-1,4-glucan-p
54	5	55.6	6	2	I48126	alpha-tubulin - Ch
55	5	55.6	6	2	B33932	Ig mu chain D regi
56	5	55.6	6	2	PT0510	T-cell receptor be
57	5	55.6	6	2	PT0518	T-cell receptor be
58	5	55.6	6	2	PT0618	T-cell receptor be
59	5	55.6	6	2	PT0662	T-cell receptor be
60	5	55.6	6	2	PT0533	T-cell receptor be
61	5	55.6	6	2	PT0568	T-cell receptor be
62	5	55.6	6	2	A41946	T-cell receptor ga
63	5	55.6	6	2	I49424	cytotoxic T-lympho
64	5	55.6	6	2	PC4127	hypothetical 6 pro
65	5	55.6	6	2	A49792	acylaminoacyl-pept
66	5	55.6	6	2	A43129	neuropeptide GNFR
67	5	55.6	6	4	I79564	hypothetical rcl3
68	5	55.6	7	2	A60224	Met-enkephalin-Arg
69	5	55.6	7	2	PH1408	Ig heavy chain V r
70	5	55.6	7	2	S19630	ribosomal protein
71	5	55.6	7	2	S16364	opacity protein P.
72	5	55.6	7	2	S18365	opacity protein P.
73	5	55.6	7	2	S55548	mcrB protein - Esc
74	5	55.6	7	2	S42407	gramicidin S synth
75	5	55.6	7	2	A28709	phosphonoacetaldeh
76	5	55.6	7	2	S38516	mablin II chain
77	5	55.6	7	2	A34818	vicilin 72K chain
78	5	55.6	7	2	PN0150	omega-gliadine 1'
79	5	55.6	7	2	S42620	aggreccan - bovine
80	5	55.6	7	2	E48394	glycoprotein compo
81	5	55.6	7	2	I48105	dihydrofolate redu
82	5	55.6	7	2	B33541	hypothetical prote
83	5	55.6	7	2	E33932	Ig mu chain D regi
84	5	55.6	7	2	PT0602	T-cell receptor be
85	5	55.6	7	2	PT0620	T-cell receptor be
86	5	55.6	7	2	PT0667	T-cell receptor be
87	5	55.6	7	2	PT0655	T-cell receptor be
88	5	55.6	7	2	PT0556	T-cell receptor be
89	5	55.6	7	2	PT0542	T-cell receptor be
90	5	55.6	7	2	PT0567	T-cell receptor be
91	5	55.6	7	2	PT0676	T-cell receptor be
92	5	55.6	7	2	PT0576	T-cell receptor be
93	5	55.6	7	2	PT0581	T-cell receptor be
94	5	55.6	7	2	PT0671	neural cell adhesi
95	5	55.6	7	2	A39690	T-cell receptor be
96	5	55.6	7	2	PH0932	major fat-globule
97	5	55.6	7	2	B48394	carnocin UI49 - Ca
98	5	55.6	7	2	A58718	NADH2 dehydrogenas
99	5	55.6	7	2	PQ0777	globulin IV alpha
100	5	55.6	7	2	S09066	protein kinase C i
101	5	55.6	7	2	A59489	choleline oxidase (E
102	5	55.6	7	2	A15398	

103	5	55.6	7	2	S45311	microcin C7 - Esch	176	5	55.6	9	2	PT0315	Ig heavy chain CRD
104	5	55.6	7	2	PC2132	FMRPamide-related	177	5	55.6	9	2	PT0326	Ig heavy chain CRD
105	5	55.6	7	2	A28340	myomodulin - Calif	178	5	55.6	9	2	C60070	gastrin - domestic
106	5	55.6	7	2	S33245	neuromodulatory pe	179	5	55.6	9	2	G56978	collagen alpha 1(I
107	5	55.6	7	2	B44787	calliFMRPamide 11	180	5	55.6	9	2	PH1591	Ig H chain V-D-J r
108	5	55.6	7	2	ECMUCR	catch-relaxing pep	181	5	55.6	9	2	S36850	Ig heavy chain V r
109	5	55.6	7	4	IS5382	hypothetical pepti	182	5	55.6	9	2	PT0670	T-cell receptor be
110	5	55.6	7	4	I56695	hypothetical L2 pr	183	5	55.6	9	2	PT0634	T-cell receptor be
111	5	55.6	7	4	S15597	orf 4 rara 5'-regi	184	5	55.6	9	2	S78420	ribosomal protein
112	5	55.6	8	1	LFSAME	probable msrA lead	185	5	55.6	9	2	S39040	lysine-conopressin
113	5	55.6	8	2	PH1407	Ig heavy chain V r	186	5	55.6	9	2	S02384	probable membrane
114	5	55.6	8	2	PC4131	hypothetical prote	187	5	55.6	9	2	B39841	orf downstream to b
115	5	55.6	8	2	T10077	hypothetical prote	188	5	55.6	9	2	B41983	dextranucrase (EC
116	5	55.6	8	2	S70727	ipgr protein - Shi	189	5	55.6	9	2	A39841	sucrose 3-glucosyl
117	5	55.6	8	2	B39745	endoglycosylcerami	190	5	55.6	9	2	A1576	xylose isomerase (
118	5	55.6	8	2	A39892	p element, P cytot	191	5	55.6	9	2	PC2021	oxytocin-related p
119	5	55.6	8	2	PT0279	Ig heavy chain CRD	192	5	55.6	9	2	B57444	neuropeptide Grb-A
120	5	55.6	8	2	PT0298	Ig heavy chain CRD	193	5	55.6	9	2	C57444	neuropeptide Grb-A
121	5	55.6	8	2	PT0323	Ig heavy chain CRD	194	5	55.6	9	2	D57444	neuropeptide Grb-A
122	5	55.6	8	2	B45800	serum albumin - do	195	5	55.6	9	2	A29477	diuretic neuropt
123	5	55.6	8	2	PH1618	Ig H chain V-D-J r	196	5	55.6	9	2	A61620	locustamyotropin I
124	5	55.6	8	2	A35768	T-cell receptor al	197	5	55.6	9	2	A44787	calliFMRPamide 1 -
125	5	55.6	8	2	PT0522	T-cell receptor be	198	5	55.6	9	2	D44787	calliFMRPamide 10
126	5	55.6	8	2	PT0639	T-cell receptor be	199	5	55.6	9	2	B41978	calliFMRPamide 2 -
127	5	55.6	8	2	PT0613	T-cell receptor be	200	5	55.6	9	2	B41978	calliFMRPamide 3 -
128	5	55.6	8	2	PT0653	T-cell receptor be	201	5	55.6	9	2	D41978	calliFMRPamide 4 -
129	5	55.6	8	2	PT0547	T-cell receptor be	202	5	55.6	9	2	B41978	calliFMRPamide 5 -
130	5	55.6	8	2	PT0691	T-cell receptor be	203	5	55.6	9	2	F41978	calliFMRPamide 6 -
131	5	55.6	8	2	B54823	olfactory receptor	204	5	55.6	9	2	G41978	calliFMRPamide 7 -
132	5	55.6	8	2	I57532	gene Th1slow prote	205	5	55.6	9	2	S06375	arginine vasotocin
133	5	55.6	8	2	A37521	R-phycocerythrin ga	206	5	55.6	9	2	B61364	vasotocin - common
134	5	55.6	8	2	S11078	glucose-6-phosphat	207	5	55.6	9	2	I50633	c-rel protein - ch
135	5	55.6	8	2	PT0030	inulinase (EC 3.2.	208	5	55.6	9	2	A54379	gene NF2 protein -
136	5	55.6	8	2	PL0162	paramyosin - north	209	5	55.6	9	2	A37027	macrophage chemota
137	5	55.6	8	2	A23967	leucopyrokinin - M	210	5	55.6	9	2	S78762	ribosomal protein
138	5	55.6	8	2	S66646	cardioacceleratory	211	5	55.6	9	2	PC7076	spectrin alpha cha
139	5	55.6	8	2	H41978	calliFMRPamide 8 -	212	5	55.6	9	2	I49406	bone gla protein -
140	5	55.6	8	2	D47393	neuropeptide calla	213	5	55.6	9	2	S39437	D-amino-acid oxida
141	5	55.6	8	2	S71919	alcohol dehydrogen	214	5	55.6	9	2	A43065	hydroxyproline-3-b
142	5	55.6	8	2	A54823	olfactory receptor	215	5	55.6	9	2	PD0027	pev-tachykinin - p
143	5	55.6	8	2	I57018	gene Cfr protein	216	5	55.6	9	2	B60246	ornitho-kinin - ch
144	5	55.6	8	2	S65381	cytochrome-c oxida	217	5	55.6	9	2	A26744	bradykinin-like pe
145	5	55.6	8	2	T13818	cytochrome oxidase	218	5	55.6	9	2	A61057	Thr-6 bradykinin -
146	5	55.6	8	2	S69165	ferredoxin a2 - Ja	219	5	55.6	9	2	A60579	bradykinin-like pe
147	5	55.6	8	2	A59028	MHC class I histoc	220	5	55.6	9	2	B45020	probable minipolyp
148	5	55.6	8	2	I64832	Ca2+-transporting	221	5	55.6	9	2	A61363	bradykinin - commo
149	5	55.6	8	2	A25836	L-serine ammonia-1	222	5	55.6	9	2	A61358	bradykinin-like pe
150	5	55.6	8	2	T48890	hypothetical prote	223	5	55.6	9	2	S15594	orf 1 rara 5'-regi
151	5	55.6	8	2	B24749	neuropeptide B - b	224	5	55.6	9	4	S15595	orf 2 rara 5'-regi
152	5	55.6	8	4	I55411	hypothetical histo	225	5	55.6	9	4	I73804	hypothetical E2 pr
153	5	55.6	9	2	B45796	dihydrolipoamide S	226	5	55.6	10	1	XAVI6B	angiotensin-conver
154	5	55.6	9	2	S59902	glutathione transf	227	5	55.6	10	1	RHSHG	gonadoliberin - pi
155	5	55.6	9	2	S13889	phosphoenolpyruvat	228	5	55.6	10	1	RHSHG	tachykinin I - mig
156	5	55.6	9	2	A28495	conopressin G - co	229	5	55.6	10	1	ECLQ1M	tachykinin III - m
157	5	55.6	9	2	B28495	conopressin S - co	230	5	55.6	10	1	ECLQ3M	tachykinin IV - mi
158	5	55.6	9	2	S35538	ribosomal protein	231	5	55.6	10	1	ECLQ4M	leucosulfakinin-II
159	5	55.6	9	2	A44873	caldesmon - rabbit	232	5	55.6	10	1	GMR0L2	cytochrome b559 co
160	5	55.6	9	2	D24180	fibrinogen beta ch	233	5	55.6	10	2	S28055	glutathione transf
161	5	55.6	9	2	C24180	fibrinogen beta ch	234	5	55.6	10	2	S71868	calpain (EC 3.4.22
162	5	55.6	9	2	E28854	fibrinopeptide B -	235	5	55.6	10	2	S39392	angiotensin I - Ja
163	5	55.6	9	2	F28854	fibrinopeptide B -	236	5	55.6	10	2	A60624	hydnin 2 - bullfro
164	5	55.6	9	2	D28854	fibrinopeptide B -	237	5	55.6	10	2	A61131	beta-neoendorphin
165	5	55.6	9	2	D58503	translation elonga	238	5	55.6	10	2	S39374	mannose receptor -
166	5	55.6	9	2	JS0302	xenopsin-related p	239	5	55.6	10	2	S66248	processing enzyme,
167	5	55.6	9	2	A60320	xenopsin-related p	240	5	55.6	10	2	S39374	hemagglutinin - in
168	5	55.6	9	2	S66607	quinoline 2-oxidor	241	5	55.6	10	2	S51912	probable methonin
169	5	55.6	9	2	C36730	hutU protein - Kle	242	5	55.6	10	2	S39517	beta-galactosidase
170	5	55.6	9	2	T31612	hypothetical prote	243	5	55.6	10	2	S59625	heat shock protein
171	5	55.6	9	2	S65433	bradykinin - horn	244	5	55.6	10	2	S70721	hypothetical prote
172	5	55.6	9	2	A53797	3',5'-cyclic-GWP p	245	5	55.6	10	2	C39191	paraspinal crystal
173	5	55.6	9	2	PT0247	Ig heavy chain CRD	246	5	55.6	10	2	S42282	seed storage prote
174	5	55.6	9	2	PT0268	Ig heavy chain CRD	247	5	55.6	10	2	C38925	
175	5	55.6	9	2	PT0299	Ig heavy chain CRD	248	5	55.6	10	2		

249	5	55.6	10	2	JN0440	peptide-N4-(N'-acet
250	5	55.6	10	2	S65432	angiotensin I - ho
251	5	55.6	10	2	PT0230	Ig heavy chain CDR
252	5	55.6	10	2	PT0245	Ig heavy chain CRD
253	5	55.6	10	2	PT0251	Ig heavy chain CRD
254	5	55.6	10	2	PT0284	Ig heavy chain CRD
255	5	55.6	10	2	PT0309	Ig heavy chain CRD
256	5	55.6	10	2	S23370	T-cell receptor al
257	5	55.6	10	2	F49033	T-cell receptor ga
258	5	55.6	10	2	B24736	inhibin beta-B cha
259	5	55.6	10	2	PH1948	matrix metalloprot
260	5	55.6	10	2	PH1633	Ig H chain V-D-J r
261	5	55.6	10	2	PH1592	Ig H chain V-D-J r
262	5	55.6	10	2	S36849	Ig heavy chain V r
263	5	55.6	10	2	I48778	small nuclear ribo
264	5	55.6	10	2	PH0807	T-cell receptor al
265	5	55.6	10	2	PT0212	T-cell receptor al
266	5	55.6	10	2	B38887	T-cell receptor ga
267	5	55.6	10	2	D54823	olfactory receptor
268	5	55.6	10	2	A55695	proteoglycan core
269	5	55.6	10	2	PH0933	T-cell receptor be
270	5	55.6	10	2	PH0894	T-cell receptor be
271	5	55.6	10	2	PH0926	T-cell receptor be
272	5	55.6	10	2	PH0923	T-cell receptor be
273	5	55.6	10	2	PH0895	T-cell receptor be
274	5	55.6	10	2	PC4374	telomeric and tetr
275	5	55.6	10	2	I52645	gene B-50 protein
276	5	55.6	10	2	D37397	hypothetical prote
277	5	55.6	10	2	JQ0943	hypothetical i.3K
278	5	55.6	10	2	S06964	hypothetical prote
279	5	55.6	10	2	S70722	65.4K GRP-binding
280	5	55.6	10	2	A60722	cryptic fibrillar p
281	5	55.6	10	2	PC4442	cytochrome c553 -
282	5	55.6	10	2	S70251	nitrogenase (NC 1-
283	5	55.6	10	2	I40032	trpE protein - Bac
284	5	55.6	10	2	A44646	neurotoxin-asfocia
285	5	55.6	10	2	I44644	neurotoxin-asfocia
286	5	55.6	10	2	F41839	ribosomal protein
287	5	55.6	10	2	A40753	aldehyde ferrédoxi
288	5	55.6	10	2	PC2044	beta-Kirilowin - M
289	5	55.6	10	2	S38305	lectin GNL2 alpha
290	5	55.6	10	2	D28027	protein P7 - curle
291	5	55.6	10	2	A27617	triase-phosphate i
292	5	55.6	10	2	PS0451	24K protein 4302 -
293	5	55.6	10	2	S19296	16K protein - pou
294	5	55.6	10	2	PK0030	triacylglycerol li
295	5	55.6	10	2	PN0165	triase-phosphate i
296	5	55.6	10	2	A58365	neuropeptide PPRFa
297	5	55.6	10	2	S30348	clotting protein -
298	5	55.6	10	2	B60656	leucosulfakinin II
299	5	55.6	10	2	A43977	FMRFamide-like pro
300	5	55.6	10	2	A42089	transcription-fact
301	5	55.6	10	2	C44787	calliFMRFamide 12
302	5	55.6	10	2	A56633	neomycinopressin -
303	5	55.6	10	2	B56899	serum heterodimer,
304	5	55.6	10	2	A90917	angiotensin precur
305	5	55.6	10	2	A90345	angiotensin precur
306	5	55.6	10	2	S26506	collagen alpha 1(V
307	5	55.6	10	2	C54823	olfactory receptor
308	5	55.6	10	2	A61218	alpha-gliadin 4Ha
309	5	55.6	10	2	B61218	alpha-gliadin 6Ha
310	5	55.6	10	2	T17054	cytochrome-c oxida
311	5	55.6	10	2	T13838	cytochrome-c oxida
312	5	55.6	10	2	T17075	cytochrome-c oxida
313	5	55.6	10	2	T13976	cytochrome-c oxida
314	5	55.6	10	2	T17057	cytochrome-c oxida
315	5	55.6	10	2	T12303	cytochrome-c oxida
316	5	55.6	10	2	T14019	cytochrome-c oxida
317	5	55.6	10	2	T17060	cytochrome-c oxida
318	5	55.6	10	2	T17063	cytochrome-c oxida
319	5	55.6	10	2	T12325	cytochrome-c oxida
320	5	55.6	10	2	T14043	cytochrome-c oxida
321	5	55.6	10	2	T14054	cytochrome-c oxida
322	5	55.6	10	2	T17066	cytochrome-c oxida
323	5	55.6	10	2	T17069	cytochrome-c oxida
324	5	55.6	10	2	T12308	cytochrome-c oxida
325	5	55.6	10	2	T17072	cytochrome-c oxida
326	5	55.6	10	2	T12312	cytochrome-c oxida
327	5	55.6	10	2	T12329	cytochrome-c oxida
328	5	55.6	10	2	T12316	cytochrome-c oxida
329	5	55.6	10	2	T14212	cytochrome-c oxida
330	5	55.6	10	2	T12321	cytochrome-c oxida
331	5	55.6	10	2	T14215	cytochrome-c oxida
332	5	55.6	10	2	T14223	cytochrome-c oxida
333	5	55.6	10	2	T14219	cytochrome-c oxida
334	5	55.6	10	2	C54226	light-harvesting p
335	5	55.6	10	2	PQ0785	NADH2 dehydrogenas
336	5	55.6	10	2	PQ0784	NADH2 dehydrogenas
337	5	55.6	10	2	S39030	lysyl-bradykinin -
338	5	55.6	10	2	S07202	phylomedulin - tw
339	5	55.6	10	2	G60589	sperm-activating p
340	5	55.6	10	2	B33710	ornithine decarbox
341	5	55.6	10	2	A32543	cardioexcitatory n
342	5	55.6	10	2	S15118	dihydrofolate redu
343	5	55.6	10	2	XASNBA	bradykinin-potenti
344	5	55.6	10	1	XAVIBH	bradykinin-potenti
345	5	55.6	10	1	ECLQ2M	tachykinin II - mi
346	5	55.6	10	1	A60654	substance P - guin
347	5	55.6	10	1	SPHO	substance P - hors
348	5	55.6	10	1	GMROL	leucosulfakinin -
349	5	55.6	10	2	S66196	alcohol dehydrogen
350	5	55.6	10	2	A33917	dihydroorotase (EC
351	5	55.6	10	2	JN0023	substance P - chic
352	5	55.6	10	2	S32575	ribosomal protein
353	5	55.6	10	2	A40693	transgelin - sheep
354	5	55.6	10	2	CS3652	RHR protein - Pse
355	5	55.6	10	2	I54193	Rhesus blood group
356	5	55.6	10	2	D58502	27K bile and gallb
357	5	55.6	10	2	JQ0395	hypothetical prote
358	5	55.6	10	2	S42587	celF protein - Esc
359	5	55.6	10	2	S33782	acetylactate synth
360	5	55.6	10	2	B43669	hypothetical prote
361	5	55.6	10	2	E60691	phycobilisome 8K 1
362	5	55.6	10	2	PC2372	58K heat shock pro
363	5	55.6	10	2	B41835	translational elonga
364	5	55.6	10	2	E41476	probable antigen 5
365	5	55.6	10	2	S19301	endo-1,4-beta-xyla
366	5	55.6	10	2	PT0081	protein QA300023 -
367	5	55.6	10	2	T06383	hypothetical prote
368	5	55.6	10	2	S78026	ribosomal protein
369	5	55.6	10	2	PN0167	ribosomal protein
370	5	55.6	10	2	A34135	DNA-binding protei
371	5	55.6	10	2	I52980	glucocorticoidase
372	5	55.6	10	2	PT0249	Ig heavy chain CRD
373	5	55.6	10	2	PT0250	Ig heavy chain CRD
374	5	55.6	10	2	PT0273	Ig heavy chain CRD
375	5	55.6	10	2	PT0287	Ig heavy chain CRD
376	5	55.6	10	2	PT0302	Ig heavy chain CRD
377	5	55.6	10	2	D56979	collagen alpha 1(I
378	5	55.6	10	2	A33571	follicstatin - bovi
379	5	55.6	10	2	PH1632	Ig H chain V-D-J r
380	5	55.6	10	2	PH1600	Ig H chain V-D-J r
381	5	55.6	10	2	PH1583	Ig H chain V-D-J r
382	5	55.6	10	2	PD0442	Ig H chain V-D-J r
383	5	55.6	10	2	PT0217	NIPSNAP2 protein -
384	5	55.6	10	2	PT0214	T-cell receptor be
385	5	55.6	10	2	PT0214	T-cell receptor be
386	5	55.6	10	2	C38887	T-cell receptor ga
387	5	55.6	10	2	B41946	T-cell receptor ga
388	5	55.6	10	2	A49037	Tcr gamma V-J regi
389	5	55.6	10	2	PD0441	translational elonga
390	5	55.6	10	2	S53436	beta-D-galactosida
391	5	55.6	10	2	S45386	low density lipopr
392	5	55.6	10	2	S78422	ribosomal protein
393	5	55.6	10	2	PH0929	T-cell receptor be
394	5	55.6	10	2	PH0938	T-cell receptor be

395	5	55.6	11	2	PH0947	T-cell receptdr be	468	12	2	A33099	163K exoantigen -
396	5	55.6	11	2	PH0903	T-cell receptor be	469	12	2	S65730	hemoglobin, extrac
397	5	55.6	11	2	PH0919	T-cell receptor be	470	12	2	C61308	hemocyanin chain 6
398	5	55.6	11	2	S60294	tubulin 2 beta-3 c	471	12	2	S10624	lipovitellin - Afr
399	5	55.6	11	2	S42449	ant1 protein - pha	472	12	2	S10626	lipovitellin - Afr
400	5	55.6	11	2	S66606	quinoline 2-oxid	473	12	2	PT0228	Ig heavy chain CDR
401	5	55.6	11	2	S04875	nifs protein - Bra	474	12	2	PT0255	Ig heavy chain CRD
402	5	55.6	11	2	A44755	20alpha-hydroxate	475	12	2	S57570	T cell receptor V-
403	5	55.6	11	2	C59151	protein-tyrosine k	476	12	2	S47360	T-cell antigen rec
404	5	55.6	11	2	S46120	6-phosphofructokin	477	12	2	S47363	T-cell antigen rec
405	5	55.6	11	2	S45698	gamma-MSH-like pro	478	12	2	PH1182	T-cell receptor al
406	5	55.6	11	2	A34662	Achatina cardio-ex	479	12	2	PH1174	T-cell receptor al
407	5	55.6	11	2	S05002	corazonin - Americ	480	12	2	PH1181	T-cell receptor al
408	5	55.6	11	2	A60656	perisulfakinin - A	481	12	2	S68402	NAD(+)-glycohydrol
409	5	55.6	11	2	S65395	chemical-sense-rel	482	12	2	A37933	Ig lambda chain J
410	5	55.6	11	2	I41978	calliWRPamide 9 -	483	12	2	C30503	Ig lambda-2b chain
411	5	55.6	11	2	S33300	probable substance	484	12	2	PH1587	Ig H chain V-D-J r
412	5	55.6	11	2	C37196	bradykinin-potenti	485	12	2	PH1605	Ig H chain V-D-J r
413	5	55.6	11	2	D37196	galbladder stone	486	12	2	PH1606	Ig H chain V-D-J r
414	5	55.6	11	2	E57789	Ile-Ser-bradykinin	487	12	2	PH1611	Ig H chain V-D-J r
415	5	55.6	11	2	S13279	amine oxidase (cop	488	12	2	PH1581	Ig H chain V-D-J r
416	5	55.6	11	2	A32428	pyridoxal kinase (	489	12	2	S25039	Ig heavy chain V r
417	5	55.6	11	2	A61483	conotoxin ausa - c	490	12	2	S68271	major urinary prot
418	5	55.6	11	2	A59146	conotoxin ausB - c	491	12	2	PH0802	T-cell receptor al
419	5	55.6	11	2	B59146	cytochrome-c oxida	492	12	2	PH0790	T-cell receptor al
420	5	55.6	11	2	T12264	cytochrome-c oxida	493	12	2	PH0746	T-cell receptor be
421	5	55.6	11	2	T12253	cytochrome-c oxida	494	12	2	A42324	cytochrome P450c27
422	5	55.6	11	2	T12244	cytochrome-c oxida	495	12	2	B39690	neural cell adhesi
423	5	55.6	11	2	T12248	cytochrome-c oxida	496	12	2	PH0930	T-cell receptor be
424	5	55.6	11	2	T17078	cytochrome-c oxida	497	12	2	PH0931	T-cell receptor be
425	5	55.6	11	2	PU0034	dextranucrase (EC	498	12	2	A61332	Na+/K+-exchanging
426	5	55.6	11	2	A48973	glucoamylase A1 (E	499	12	2	S21163	NAD ADP-ribosyltra
427	5	55.6	11	2	S23308	substance P - rain	500	12	2	S29859	gene p10 protein -
428	5	55.6	11	2	S23306	substance P - Atla	501	12	2	S71034	poB protein - Sal
429	5	55.6	11	2	S07207	Crinia-angiotensin	502	12	2	S29479	hypothetical prote
430	5	55.6	11	2	D61033	ranatachykinin D -	503	12	2	T44420	hypothetical prote
431	5	55.6	11	2	A26930	ermG leader peptid	504	12	2	JQ2308	hypothetical 1.4K
432	5	55.6	11	2	B26744	megascoliakinin -	505	12	2	JQ2318	hypothetical 1.4K
433	5	55.6	11	2	A61365	phyllokinin - Rohd	506	12	2	S17540	agglutinin-III lec
434	5	55.6	11	2	A61033	ranatachykinin A -	507	12	2	PN0160	ribosomal protein
435	5	55.6	11	2	A57458	gene Gax protein -	508	12	2	T46656	hypothetical prote
436	5	55.6	11	2	S41909	hypothetical prote	509	12	2	S16204	hypothetical prote
437	5	55.6	11	4	PC2124	aminotransferase c	510	12	2	A56878	light yellow cell
438	5	55.6	11	4	I52708	ELAV-like neuronal	511	12	2	E44787	light yellow cell
439	5	55.6	11	4	S19015	hypothetical prote	512	12	2	A33900	calliWRamide 1 -
440	5	55.6	11	4	I54081	retinoic acid rece	513	12	2	A49637	hydrlin 1 - African
441	5	55.6	12	1	A43975	locustamytotropin -	514	12	2	A34858	MHC class II histo
442	5	55.6	12	1	A53709	alpha-conotoxin Im	515	12	2	S43170	kinasin light chai
443	5	55.6	12	1	LFECPE	pyrE leader peptid	516	12	2	A60528	insulin-like growt
444	5	55.6	12	1	A29169	phospholipase A2 (	517	12	2	I46922	gene Bata protein
445	5	55.6	12	2	I64829	gene HEXA protein	518	12	2	PQ0776	inhibitory diffusl
446	5	55.6	12	2	A49215	urease (EC 3.9.1.5	519	12	2	S07436	NADH2 dehydrogenas
447	5	55.6	12	2	C36201	1-aminocyclopropan	520	12	2	S07436	tachykinin - Afric
448	5	55.6	12	2	PH1675	Ig heavy chain V r	521	12	2	A61360	vespakinin M - hor
449	5	55.6	12	2	S25056	Ig heavy chain - m	522	12	2	A61359	vespakinin X - hor
450	5	55.6	12	2	S26557	T-cell receptor be	523	12	2	A44874	probooscipedia - fr
451	5	55.6	12	2	S26556	T-cell receptor be	524	12	4	JX0315	aminotransferase c
452	5	55.6	12	2	S26554	T-cell receptor be	525	12	4	PC2122	aminotransferase c
453	5	55.6	12	2	S26555	T-cell receptor be	526	12	4	PC2123	frame shifted cyta
454	5	55.6	12	2	S26541	T-cell receptor be	527	12	4	S49073	angiotensin-conver
455	5	55.6	12	2	A54315	entactin/nidogen -	528	13	1	XAV19B	angiotensin - bovi
456	5	55.6	12	2	S36902	Em protein - wheat	529	13	1	UNBO	neurotensin - bovi
457	5	55.6	12	2	S01222	translation elonga	530	13	1	UNOPBT	neurotensin - brus
458	5	55.6	12	2	S15815	protoporphyrinogen	531	13	1	MTCMAD	melanotropin alpha
459	5	55.6	12	2	S65629	proton-translocati	532	13	1	JTUG3	melanotropin alpha
460	5	55.6	12	2	S69123	binR protein - Sta	533	13	1	MTUG3	trimerogen a-13 -
461	5	55.6	12	2	S34447	sucrose-6-phosphat	534	13	2	A39229	glutathione peroxi
462	5	55.6	12	2	A40763	hypothetical prote	535	13	2	A28505	neurotensin [valid
463	5	55.6	12	2	D28551	hypothetical prote	536	13	2	A53608	neurotensin - comm
464	5	55.6	12	2	S49547	hypothetical prote	537	13	2	A53608	neurotensin - guin
465	5	55.6	12	2	S36899	ribosomal protein	538	13	2	A32734	enkephalin precurs
466	5	55.6	12	2	S67528	napin - rape (frag	539	13	2	PH1676	Ig heavy chain V r
467	5	55.6	12	2	S18722	matk protein - bee	540	13	2	S78519	ribosomal protein

541	5	55.6	13	2	S48210	collagen alpha 1(V	614	13	2	AB0764	his operon leader
542	5	55.6	13	2	E39778	lactose phosphotra	615	13	2	I54984	aeg-46.5 protein -
543	5	55.6	13	2	PN0122	O1L protein - vacc	616	13	2	PC1008	40K extracellular
544	5	55.6	13	2	A60458	protocatechuate 3,	617	13	2	S36668	hypothetical prote
545	5	55.6	13	2	T08533	hypothetical prote	618	13	2	PC2369	unidentified 85K p
546	5	55.6	13	2	S22995	hypothetical prote	619	13	2	S20578	ribosomal protein
547	5	55.6	13	2	S12388	argA protein - Sal	620	13	2	S09733	photosystem I prot
548	5	55.6	13	2	PC2371	probable endopepti	621	13	2	JQ2309	hypothetical 1.6K
549	5	55.6	13	2	S36887	ribosomal protein	622	13	2	JQ2319	hypothetical 1.6K
550	5	55.6	13	2	S23103	erythronolide synt	623	13	2	PN0168	phosphopyruvate hy
551	5	55.6	13	2	S09716	2S albumin large c	624	13	2	S32471	lymaADFamide 1 - g
552	5	55.6	13	2	H44957	protein P18 - Comm	625	13	2	S32472	lymaADFamide 2 - g
553	5	55.6	13	2	PA0089	protein QP200053 -	626	13	2	S32473	lymaADFamide 3 - g
554	5	55.6	13	2	E60396	potassium channel	627	13	2	S32474	lymaADFamide 4 - g
555	5	55.6	13	2	PS0443	antigen 7H8/2 - ma	628	13	2	S32475	lymaADFamide 5 - g
556	5	55.6	13	2	S52356	hypothetical prote	629	13	2	B61620	locustamyotropin I
557	5	55.6	13	2	PT0256	Ig heavy chain CRD	630	13	2	I84603	deoxynucleotidyltr
558	5	55.6	13	2	PT0293	Ig heavy chain CRD	631	13	2	S23638	Ig kappa chain J s
559	5	55.6	13	2	PT0304	Ig heavy chain CRD	632	13	2	S23638	urinary tract ston
560	5	55.6	13	2	PT0305	Ig heavy chain CRD	633	13	2	S03879	6-phosphofructokin
561	5	55.6	13	2	PT0331	Ig heavy chain CRD	634	13	2	S32551	glutathione transf
562	5	55.6	13	2	PH1316	Ig heavy chain DJ	635	13	2	H33932	Ig kappa chain J r
563	5	55.6	13	2	PH1309	Ig heavy chain DJ	636	13	2	A33933	Ig kappa chain J r
564	5	55.6	13	2	S23640	Ig kappa chain J s	637	13	2	PC4391	cysteine proteinas
565	5	55.6	13	2	S70441	pancreatic elatase	638	13	2	G83988	hypothetical prote
566	5	55.6	13	2	S47356	T-cell antigen rec	639	13	2	H85575	hypothetical prote
567	5	55.6	13	2	S47357	T-cell antigen rec	640	13	2	S01043	glutamate-ammonia
568	5	55.6	13	2	S47358	T-cell antigen rec	641	13	2	I50173	alpha-2 collagen -
569	5	55.6	13	2	S47359	T-cell antigen rec	642	13	2	JH0460	corticostatic pept
570	5	55.6	13	2	S47365	T-cell antigen rec	643	13	2	JZVHP1	crabrolin - Europe
571	5	55.6	13	2	S47368	T-cell antigen rec	644	13	2	A61361	bradykinin-like pe
572	5	55.6	13	2	S47371	T-cell antigen rec	645	13	4	I70075	glycophorin B (mis
573	5	55.6	13	2	S47372	T-cell antigen rec	646	13	1	NYPG14	hypothalamic tetra
574	5	55.6	13	2	S47373	T-cell antigen rec	647	13	1	NTKN1M	alpha-conotoxin MI
575	5	55.6	13	2	S47376	T-cell antigen rec	648	13	1	LFEBWC	trp operon leader
576	5	55.6	13	2	S47377	T-cell antigen rec	649	13	1	LFEBWT	trp operon leader
577	5	55.6	13	2	S47380	T-cell antigen rec	650	13	1	LFECW	trp operon leader
578	5	55.6	13	2	S47381	T-cell antigen rec	651	13	1	LFECFS	phesT operon leade
579	5	55.6	13	2	S47382	T-cell antigen rec	652	13	1	BSTD	bombesin - fire-be
580	5	55.6	13	2	S47384	T-cell antigen rec	653	13	2	A47421	leukotriene B-4 12
581	5	55.6	13	2	S47385	T-cell antigen rec	654	13	2	A33798	D-amino-acid oxida
582	5	55.6	13	2	S47388	T-cell antigen rec	655	13	2	C40944	hypothetical prote
583	5	55.6	13	2	S47389	T-cell antigen rec	656	13	2	A01250	angiotensin precu
584	5	55.6	13	2	S47390	T-cell antigen rec	657	13	2	B61309	lutropin beta chai
585	5	55.6	13	2	S47392	T-cell antigen rec	658	13	2	A58963	alpha-conotoxin Cn
586	5	55.6	13	2	S47374	T-cell antigen rec	659	13	2	PH1677	Ig heavy chain V r
587	5	55.6	13	2	S23372	T-cell receptor al	660	13	2	PH1705	Ig heavy chain V r
588	5	55.6	13	2	S61798	T-cell-specific tr	661	13	2	I51432	histone H4-1 precu
589	5	55.6	13	2	B47415	mannose-1-phosphat	662	13	2	A49018	myosin heavy chain
590	5	55.6	13	2	S10562	zona pellucida-bin	663	13	2	T46634	acyl carrier prote
591	5	55.6	13	2	A39836	aggreccan - bovine	664	13	2	S50900	chlorophyll a/b-bi
592	5	55.6	13	2	B56964	dipeptidyl-peptida	665	13	2	B90858	trp operon leader
593	5	55.6	13	2	C53275	Ig kappa-1 chain J	666	13	2	B85761	trp operon leader
594	5	55.6	13	2	E53275	Ig kappa-1 chain J	667	13	2	F90931	phesT operon leade
595	5	55.6	13	2	PH1636	Ig H chain V-D-J r	668	13	2	B85780	phesT operon leade
596	5	55.6	13	2	PH1620	Ig H chain V-D-J r	669	13	2	S60353	amylopullulanase -
597	5	55.6	13	2	PH1593	Ig H chain V-D-J r	670	13	2	A58503	kidney and bladder
598	5	55.6	13	2	PH1595	Ig H chain V-D-J r	671	13	2	PL0142	carbon-monoxide de
599	5	55.6	13	2	PH1596	Ig H chain V-D-J r	672	13	2	PS0371	hypothetical prote
600	5	55.6	13	2	PH1585	Ig H chain V-D-J r	673	13	2	S36892	ribosomal protein
601	5	55.6	13	2	B26406	Ig kappa chain J r	674	13	2	S29789	hypothetical prote
602	5	55.6	13	2	S22761	Ig lambda-2 chain	675	13	2	PA0111	protein QAI00054 -
603	5	55.6	13	2	PH0788	T-cell receptor al	676	13	2	S09721	2S albumin small c
604	5	55.6	13	2	PH0787	T-cell receptor al	677	13	2	PN0151	omega-gliadin 2'
605	5	55.6	13	2	PH0799	T-cell receptor al	678	13	2	PA0104	protein QP200070 -
606	5	55.6	13	2	PH0805	T-cell receptor al	679	13	2	B34135	DNA-binding protei
607	5	55.6	13	2	A47630	Ig kappa chain J r	680	13	2	PC1215	homeotic protein E
608	5	55.6	13	2	C47630	Ig kappa chain J r	681	13	2	S23376	collagen alpha cha
609	5	55.6	13	2	B47630	Ig kappa chain J r	682	13	2	PL0152	metal-binding prot
610	5	55.6	13	2	D47630	Ig kappa chain J r	683	13	2	E61308	hemocyanin chain 3
611	5	55.6	13	2	D47630	Ig kappa chain J r	684	13	2	S38307	DEB-A protein - fr
612	5	55.6	13	2	I51905	collecting duct wa	685	13	2	S11074	alcohol dehydrogen
613	5	55.6	13	2	S54344	glyceraldehyde-3-p	686	13	2	I54284	C1-inhibitor - hum

687	5	55.6	14	2	I64815	carbonic anhydrase	760	5	55.6	14	2	S36678	dodecenoyl-CoA Del
688	5	55.6	14	2	PT0223	Ig heavy chain CDR	761	5	55.6	14	2	S68095	calcium-binding pr
689	5	55.6	14	2	PT0232	Ig heavy chain CRD	762	5	55.6	14	2	H83778	hypothetical prote
690	5	55.6	14	2	PT0252	Ig heavy chain CRD	763	5	55.6	14	2	AF0296	phenylalanyl-tRNA
691	5	55.6	14	2	PT0254	Ig heavy chain CRD	764	5	55.6	14	2	S29878	Na+/K+-exchanging
692	5	55.6	14	2	PT0294	Ig heavy chain CRD	765	5	55.6	14	2	S27140	hypothetical prote
693	5	55.6	14	2	PH1347	Ig heavy chain DJ	766	5	55.6	14	2	JH0328	probursin tetradec
694	5	55.6	14	2	PH1327	Ig heavy chain DJ	767	5	55.6	14	2	A42473	ermK leader peptid
695	5	55.6	14	2	PH1332	Ig heavy chain DJ	768	5	55.6	14	2	A44515	Trp EG leader pept
696	5	55.6	14	2	PH1332	Ig heavy chain DJ	769	5	55.6	14	2	A61362	bradykinin-like pe
697	5	55.6	14	2	PH1311	Ig heavy chain DJ	770	5	55.6	14	2	A32654	fibrinogen peptide A -
698	5	55.6	14	2	PH1321	Ig heavy chain DJ	771	5	55.6	14	4	I52618	hemoglobin beta ch
699	5	55.6	14	2	PH1305	Ig heavy chain DJ	772	5	55.6	14	4	S00843	hypothetical prote
700	5	55.6	14	2	PH1306	Ig heavy chain DJ	773	5	55.6	15	1	NTKNAG	alpha-conotoxin GI
701	5	55.6	14	2	S23369	Ig kappa chain J s	774	5	55.6	15	1	LFTWL	leu leader peptide
702	5	55.6	14	2	PH1763	T cell receptor al	775	5	55.6	15	2	PA0041	plastoquinol-plasc
703	5	55.6	14	2	PH1767	T cell receptor al	776	5	55.6	15	2	S42741	ubiquinol-cytochro
704	5	55.6	14	2	S57569	T cell receptor V-	777	5	55.6	15	2	B61243	dimethylalanine mo
705	5	55.6	14	2	S23369	T-cell receptor al	778	5	55.6	15	2	S14749	3-dehydroquinase -
706	5	55.6	14	2	PH0135	very late antigen-	779	5	55.6	15	2	S21241	oligo-1,6-glucosid
707	5	55.6	14	2	B28018	spermadhesin AWN h	780	5	55.6	15	2	S4159	leukocyte elastase
708	5	55.6	14	2	F48394	glycoprotein compo	781	5	55.6	15	2	PC2215	basic proteinase I
709	5	55.6	14	2	A43847	Ig mu chain V regi	782	5	55.6	15	2	A47146	topoisomerase I -
710	5	55.6	14	2	C44823	synaptosomal-assoc	783	5	55.6	15	2	A60834	angiotensin I prec
711	5	55.6	14	2	A61032	croponin 1, cardia	784	5	55.6	15	2	I49420	placental lactogen
712	5	55.6	14	2	JS0272	hypothetical 1.5K	785	5	55.6	15	2	S46525	T-cell receptor al
713	5	55.6	14	2	PH1625	Ig H chain V-D-J r	786	5	55.6	15	2	S26518	T-cell receptor al
714	5	55.6	14	2	PH1626	Ig H chain V-D-J r	787	5	55.6	15	2	S26517	T-cell receptor al
715	5	55.6	14	2	PH1627	Ig H chain V-D-J r	788	5	55.6	15	2	A38304	heterogeneous ribo
716	5	55.6	14	2	PH1628	Ig H chain V-D-J r	789	5	55.6	15	2	S36888	ribosomal protein
717	5	55.6	14	2	PH1614	Ig H chain V-D-J r	790	5	55.6	15	2	I58116	Dp116 - human
718	5	55.6	14	2	PH1639	Ig H chain V-D-J r	791	5	55.6	15	2	I46512	troponin - rabbit
719	5	55.6	14	2	PH1617	Ig H chain V-D-J r	792	5	55.6	15	2	I49407	placental calcium-
720	5	55.6	14	2	PH1623	Ig H chain V-D-J r	793	5	55.6	15	2	I29501	fibrinopeptide A -
721	5	55.6	14	2	PH1586	Ig H chain V-D-J r	794	5	55.6	15	2	P29501	fibrinopeptide A -
722	5	55.6	14	2	PH1594	Ig H chain V-D-J r	795	5	55.6	15	2	JP0101	fibrinogen alpha c
723	5	55.6	14	2	PH1597	Ig H chain V-D-J r	796	5	55.6	15	2	PQ0017	terminal protein -
724	5	55.6	14	2	PH1608	Ig H chain V-D-J r	797	5	55.6	15	2	S59492	formate dehydrogen
725	5	55.6	14	2	PH0792	T-cell receptor al	798	5	55.6	15	2	UN7030	hypothetical 1.7K
726	5	55.6	14	2	PH0804	T-cell receptor al	799	5	55.6	15	2	C48401	ribosomal protein
727	5	55.6	14	2	PT0210	T-cell receptor be	800	5	55.6	15	2	C41383	32K variable histo
728	5	55.6	14	2	PH0765	T-cell receptor be	801	5	55.6	15	2	S29386	nigerythrins - Desu
729	5	55.6	14	2	PH0762	T-cell receptor be	802	5	55.6	15	2	PA58786	pimeloyl-CoA synth
730	5	55.6	14	2	PH0755	T-cell receptor de	803	5	55.6	15	2	PA0018	photosystem I 9K p
731	5	55.6	14	2	D35141	T-cell receptor de	804	5	55.6	15	2	PA0029	protein QA100012 -
732	5	55.6	14	2	C35141	T-cell receptor de	805	5	55.6	15	2	PA0020	protein QA300050 -
733	5	55.6	14	2	E35141	T-cell receptor de	806	5	55.6	15	2	PA0024	stylar glycoprotei
734	5	55.6	14	2	S65392	cytochrome-c oxida	807	5	55.6	15	2	PQ0024	omega-gliadine 3 -
735	5	55.6	14	2	PH0945	T-cell receptor be	808	5	55.6	15	2	PQ0193	15K protein 5106 -
736	5	55.6	14	2	PH0915	T-cell receptor be	809	5	55.6	15	2	PS0251	20K protein 5403 -
737	5	55.6	14	2	C48394	major fat-globule	810	5	55.6	15	2	PS0208	alpha-globulin - r
738	5	55.6	14	2	G33160	H+-transporting tw	811	5	55.6	15	2	PC4268	unidentified QR310
739	5	55.6	14	2	A54370	inorganic diphosph	812	5	55.6	15	2	PC4269	adenylate isopente
740	5	55.6	14	2	AG0705	phenylalanyl-tRNA	813	5	55.6	15	2	PA0057	fructose-bisphosph
741	5	55.6	14	2	PH1471	T-cell receptor be	814	5	55.6	15	2	PA0076	protein QF200016 -
742	5	55.6	14	2	I56388	Km(r) protein - Es	815	5	55.6	15	2	PA0051	protein QF200037 -
743	5	55.6	14	2	A41589	25K elastin-bindin	816	5	55.6	15	2	PA0060	ICL3 protein - Par
744	5	55.6	14	2	S58862	botulinum neurotox	817	5	55.6	15	2	S71300	protein 425 - Cali
745	5	55.6	14	2	S58866	botulinum neurotox	818	5	55.6	15	2	G60977	T-cell receptor be
746	5	55.6	14	2	PA0045	porin por1 - Arabi	819	5	55.6	15	2	C36198	hypothetical prote
747	5	55.6	14	2	PT0026	calotropin DI - mu	820	5	55.6	15	2	B32800	Ig heavy chain CDR
748	5	55.6	14	2	PN0147	omega-gliadine 1 a	821	5	55.6	15	2	PT0222	Ig heavy chain DJ
749	5	55.6	14	2	A61306	ribonuclease M (EC	822	5	55.6	15	2	PH1314	Ig heavy chain DJ
750	5	55.6	14	2	S45655	cathepsin L (EC 3.	823	5	55.6	15	2	PH1342	Ig heavy chain DJ
751	5	55.6	14	2	A61308	hemocyanin chain 2	824	5	55.6	15	2	PH1329	Ig heavy chain DJ
752	5	55.6	14	2	G61308	hemocyanin chain 3	825	5	55.6	15	2	PH1310	Ig heavy chain DJ
753	5	55.6	14	2	D61308	hemocyanin chain 5	826	5	55.6	15	2	PH1788	T cell receptor al
754	5	55.6	14	2	A48389	leirutoxin III -	827	5	55.6	15	2	S57577	T cell receptor V-
755	5	55.6	14	2	A56632	neogulfakinin-II -	828	5	55.6	15	2	S57577	T-cell receptor V-
756	5	55.6	14	2	S12904	protein kinase (EC	829	5	55.6	15	2	G41299	T-cell receptor be
757	5	55.6	14	2	PL0040	glycogen phosphory	830	5	55.6	15	2	PH0136	T-cell receptor be
758	5	55.6	14	2	JH0516	insulin-like growt	831	5	55.6	15	2	G49255	T-cell receptor be
759	5	55.6	14	2	D45474	thrombospondin 2 -	832	5	55.6	15	2		



833	5	55.6	15	2	C24687	T-cell receptor be	906	15	2	PS0450	23K protein 4307 -
834	5	55.6	15	2	D28587	T-cell receptor be	907	15	2	PS0212	29K protein 4228 -
835	5	55.6	15	2	F28587	T-cell receptor be	908	15	2	PA0064	cytochrome C 1 - f
836	5	55.6	15	2	S51735	T-cell receptor be	909	15	2	PA0087	cytochrome c2 - fu
837	5	55.6	15	2	B49655	T-cell receptor be	910	15	2	PA0075	fructose-bisphosph
838	5	55.6	15	2	A61247	urogenital tumor m	911	15	2	PA0102	fructose-bisphosph
839	5	55.6	15	2	S66215	carilage oligomer	912	15	2	PA0062	fumarate hydratase
840	5	55.6	15	2	E56978	collagen alpha 2(X	913	15	2	A49177	22K protein pl, mi
841	5	55.6	15	2	B33527	fructose-2,6-bisph	914	15	2	A08416	lombrocin kinase
842	5	55.6	15	2	PQ0074	T-cell receptor be	915	15	2	A08416	allotostat in - tob
843	5	55.6	15	2	S30608	translation elonga	916	15	2	A36527	juvenile-hormone e
844	5	55.6	15	2	F44823	synaptosomal-assoc	917	15	2	PT0090	alpha-glucosidase
845	5	55.6	15	2	I78838	flt3 ligand isoform	918	15	2	I50503	agrin - electric r
846	5	55.6	15	2	A27504	histone H2A - mous	919	15	2	A26212	carboxypeptidase B
847	5	55.6	15	2	PH1613	Ig H chain V-D-J r	920	15	2	A32921	beaded-chain filam
848	5	55.6	15	2	PH1616	Ig H chain V-D-J r	921	15	2	A32971	major immunophilin
849	5	55.6	15	2	PH1590	Ig H chain V-D-J r	922	15	2	A34980	heparin-binding le
850	5	55.6	15	2	PH1610	Ig H chain V-D-J r	923	15	2	B45115	pepidylprolyl iso
851	5	55.6	15	2	PH1582	Ig H chain V-D-J r	924	15	2	B45133	casein kinase II (
852	5	55.6	15	2	PH1378	T antigen variant	925	15	2	A35417	28K serine protein
853	5	55.6	15	2	PH1377	T antigen variant	926	15	2	B45474	thrombospondin 2 -
854	5	55.6	15	2	PH0775	T-cell receptor al	927	15	2	A53085	lipid transfer pro
855	5	55.6	15	2	PH0779	T-cell receptor al	928	15	2	I46909	voltage-dependent
856	5	55.6	15	2	PH0780	T-cell receptor al	929	15	2	S62675	collagen type I -
857	5	55.6	15	2	PH0806	T-cell receptor al	930	15	2	PT0096	pyruvate dehydrog
858	5	55.6	15	2	PH0797	T-cell receptor al	931	15	2	PT0094	succinate dehydrog
859	5	55.6	15	2	PH0760	T-cell receptor be	932	15	2	I67525	CD33 antigen homol
860	5	55.6	15	2	PH0772	T-cell receptor be	933	15	2	A1902	bone acidic glycop
861	5	55.6	15	2	PH0770	T-cell receptor be	934	15	2	A56963	acid phosphatase (
862	5	55.6	15	2	PH0764	T-cell receptor be	935	15	2	S23175	D-galactose-bindin
863	5	55.6	15	2	B35141	T-cell receptor de	936	15	2	PQ0780	NADH2 dehydrogenas
864	5	55.6	15	2	G35141	T-cell receptor de	937	15	2	S04586	NADH2 dehydrogenas
865	5	55.6	15	2	A35141	T-cell receptor de	938	15	2	PQ0025	ubiquinol-cytochro
866	5	55.6	15	2	B49037	TcR delta chain V-	939	15	2	A28497	neutensin-relate
867	5	55.6	15	2	S71306	heat shock protein	940	15	2	D48648	hypothetical leade
868	5	55.6	15	2	C31409	protein 425 - rat	941	15	2	I40665	ILVEn leader pepti
869	5	55.6	15	2	S21293	RIP protein - rat	942	15	2	JT0610	leukocyte chemoatt
870	5	55.6	15	2	A45096	thyrotropin-releas	943	15	4	I38335	hypothetical TEL/M
871	5	55.6	15	2	PL0109	complement factor	944	15	4	I52698	hypothetical THRAl
872	5	55.6	15	2	S72432	epoxypropan isomer	945	15	4	I38336	hypothetical TEL/M
873	5	55.6	15	2	PT0091	H+-transporting tw	946	15	1	A49761	locustapyrokinin -
874	5	55.6	15	2	PT0095	H+-transporting tw	947	15	1	MTDFBS	melanotropin beta
875	5	55.6	15	2	S71396	dihydropyrimidine	948	15	1	LFECB	his operon leader
876	5	55.6	15	2	PT0092	NADH2 dehydrogenas	949	15	2	B42324	cytochrome P450c27
877	5	55.6	15	2	AF0832	phe leader peptide	950	15	2	S03405	hydrogenase (EC 1.
878	5	55.6	15	2	C56979	collagen alpha 1(I	951	15	2	C45133	casein kinase II (
879	5	55.6	15	2	H56978	collagen alpha 1(X	952	15	2	S65520	phospholipase A2 (
880	5	55.6	15	2	PH1455	T-cell receptor al	953	15	2	S10807	protein kinase C i
881	5	55.6	15	2	PH1443	T-cell receptor al	954	15	2	S10809	protein kinase C i
882	5	55.6	15	2	S02381	probable membrane	955	15	2	F44908	chitinase (EC 3.2.
883	5	55.6	15	2	S02381	probable membrane	956	15	2	A60551	leukocyte elastase
884	5	55.6	15	2	A60929	dichloromethane de	957	15	2	A44413	proteasome endopep
885	5	55.6	15	2	PN0629	integration host f	958	15	2	A59042	alpha-conotoxin Ep
886	5	55.6	15	2	S67918	serine proteinase	959	15	2	JH0517	insulin-like growt
887	5	55.6	15	2	PL0143	carbon-monoxide de	960	15	2	D49021	Ig heavy chain J7
888	5	55.6	15	2	B60929	dichloromethane de	961	15	2	A28144	ribosomal protein
889	5	55.6	15	2	S21411	modulation protein	962	15	2	A27803	myosin light chain
890	5	55.6	15	2	A41338	isocitrate lyase (	963	15	2	A29501	ribosomal protein A -
891	5	55.6	15	2	S71920	proteinase ECP 32	964	15	2	B24180	fibrinogen alpha c
892	5	55.6	15	2	A35389	urease (EC 3.5.1.5	965	15	2	A24180	fibrinogen alpha c
893	5	55.6	15	2	T03000	GTP-binding protei	966	15	2	B28854	fibrinogen alpha c
894	5	55.6	15	2	S33781	acetylactate synth	967	15	2	C28854	fibrinogen alpha c
895	5	55.6	15	2	C43334	orf3 3' to aadR -	968	15	2	A28854	fibrinogen alpha c
896	5	55.6	15	2	B41868	hypothetical prote	969	15	2	G29501	fibrinogen alpha c
897	5	55.6	15	2	AL7340	ribonucleoside-dip	970	15	2	H29501	fibrinogen alpha c
898	5	55.6	15	2	S36891	ribosomal protein	971	15	2	CQR	fibrinogen alpha c
899	5	55.6	15	2	S36896	ribosomal protein	972	15	2	B25979	fibrinogen alpha c
900	5	55.6	15	2	S36889	ribosomal protein	973	15	2	A45133	fibrinogen alpha c
901	5	55.6	15	2	S59489	steroid monooxygen	974	15	2	I78533	gene agouti protei
902	5	55.6	15	2	A06634	orf19 3' of eryK -	975	15	2	G45681	orf 61.1 - phage T
903	5	55.6	15	2	T09463	ribosomal protein	976	15	2	A39109	hypothetical prote
904	5	55.6	15	2	S62641	porphobilinogen sy	977	15	2	S51057	ribosomal protein
905	5	55.6	15	2	A56970	GLYMA1 - soybean (	978	15	2	S22677	heat shock protein

979 5 55.6 16 2 A53337 regulatory protein  
980 5 55.6 16 2 T37075 hypothetical prote  
981 5 55.6 16 2 S51610 hypothetical prote  
982 5 55.6 16 2 A34053 deoxyribonuclease  
983 5 55.6 16 2 B35491 retinol-binding pr  
984 5 55.6 16 2 PT0224 Ig heavy chain CDR  
985 5 55.6 16 2 PT0282 Ig heavy chain CDR  
986 5 55.6 16 2 PT0296 Ig heavy chain CDR  
987 5 55.6 16 2 PH1346 Ig heavy chain DJ  
988 5 55.6 16 2 PH1351 Ig heavy chain DJ  
989 5 55.6 16 2 PH1317 Ig heavy chain DJ  
990 5 55.6 16 2 PH1302 Ig heavy chain DJ  
991 5 55.6 16 2 PH1778 T cell receptor al  
992 5 55.6 16 2 PH1781 T cell receptor al  
993 5 55.6 16 2 PH1782 T cell receptor al  
994 5 55.6 16 2 PH1790 T cell receptor al  
995 5 55.6 16 2 PH1771 T cell receptor be  
996 5 55.6 16 2 S57517 T-cell receptor al  
997 5 55.6 16 2 F41299 T-cell receptor be  
998 5 55.6 16 2 PH0137 T-cell receptor be  
999 5 55.6 16 2 G49039 T-cell receptor be  
1000 5 55.6 16 2 A28587

## ALIGNMENTS

RESULT 1  
A22565  
R-phycoerythrin alpha-1 chain - red alga (Gastrocloonium coulteri) (fragment)  
C:Species: Gastrocloonium coulteri  
C>Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
C:Accession: A22565  
R:Klotz, A.V.; Glazer, A.N.  
J. Biol. Chem. 260, 4856-4863, 1985  
A:Title: Characterization of the bilin attachment sites in R-phycoerythrin.  
A:Reference number: A22565; MUID:85182601; PMID:3886644  
A:Accession: A22565  
A:Molecule type: protein  
A:Residues: 1-3 <KLO>  
A:Cross-references: UNIPARC:UPI000017CE9A

Query Match 55.6%; Score 5; DB 3; Length 3;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1  
|  
Db 3 R 3

RESULT 2  
PQ0010  
angiotensin-converting enzyme inhibitor (FLP-3) - common fig  
N:Alternate names: ficus latex peptide 3  
C:Species: Ficus carica (common fig)  
C>Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
C:Accession: PQ0010  
R:Maruyama, S.; Miyoshi, S.; Tanaka, H.  
Agric. Biol. Chem. 53, 2763-2767, 1989  
A:Title: Angiotensin I-converting enzyme inhibitors derived from Ficus carica.  
A:Reference number: PQ0008  
A:Accession: PQ0010  
A:Molecule type: protein  
A:Residues: 1-3 <MAR>  
A:Cross-references: UNIPARC:UPI000011E971  
A:Experimental source: latex  
C:Keywords: angiotensin-converting enzyme inhibitor

Query Match 55.6%; Score 5; DB 3; Length 3;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1  
|  
Db 3 R 3

RESULT 3  
ECKAA

antho-RFamide neuropeptide - sea anemone (Anthopleura elegantissima)  
C:Species: Anthopleura elegantissima  
C>Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 09-Jul-2004  
C:Accession: A26666  
R:Grimmelikhuizen, C.J.P.; Graff, D.  
Proc. Natl. Acad. Sci. U.S.A. 83, 9817-9821, 1986  
A:Title: Isolation of <Glu-Gly-Arg-Phe-NH2 (Antho-RFamide), a neuropeptide from sea anemone.  
A:Reference number: A26666; MUID:87092339; PMID:2879288  
A:Accession: A26666

A:Molecule type: protein  
A:Residues: 1-4 <GRI>  
A:Cross-references: UNIPROT:P10419; UNIPARC:UPI00001733AF  
C:Comment: The function of this peptide is not known but it could act as a transmitter.  
C:Comment: Synthetic and natural peptides had identical properties.  
C:Superfamily: RFamide neuropeptide  
C:Keywords: amidated carboxyl end; neuropeptide; pyroglutamic acid  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F:4/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 55.6%; Score 5; DB 1; Length 4;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1  
|  
Db 3 R 3

RESULT 4  
A02147

phagocytosis-stimulating peptide (tuftsin) - human  
C:Species: Homo sapiens (man)  
C>Date: 31-Mar-1991 #sequence\_revision 31-Mar-1991 #text\_change 09-Jul-2004  
C:Accession: A02147  
R:Nishioka, K.; Constantopoulos, A.; Satoh, P.S.; Najjar, V.A.  
Biochem. Biophys. Res. Commun. 47, 172-179, 1972  
A:Title: The characteristics, isolation and synthesis of the phagocytosis stimulating peptide  
A:Reference number: A02147; MUID:72187087; PMID:4112769  
A:Accession: A02147

A:Molecule type: protein  
A:Residues: 1-4 <NIS>  
A:Cross-references: UNIPROT:P01858; UNIPARC:UPI00001377C2  
A:Note: a peptide having the same structure, physical properties, and biological activities  
R:Fidalgo, B.V.; Najjar, V.A.  
Biochemistry 6, 3386-3392, 1967  
A:Reference number: A37502; MUID:68091045; PMID:4169272  
C:Comment: annotation; immunoglobulin class  
C:Comment: An IGG (called leucokinin) binds reversibly to the cell membrane of neutrophils.  
n is essential for maximum stimulation of the phagocytic activity of neutrophils.  
C:Superfamily: immunoglobulin C region; immunoglobulin homology

Query Match 55.6%; Score 5; DB 2; Length 4;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1  
|  
Db 4 R 4

RESULT 5  
D41654

hypothetical protein (sodC 5' region) - Haemophilus parainfluenzae (fragment)  
C:Species: Haemophilus parainfluenzae  
C>Date: 12-Jun-1992 #sequence\_revision 12-Jun-1992 #text\_change 24-Feb-1995  
C:Accession: D41654

R;Kroll, J.S.; Langford, P.R.; Loynds, B.M.  
J. Bacteriol. 173, 7449-7457, 1991  
A:Title: Copper-zinc superoxide dismutase of *Haemophilus influenzae* and *Haemophilus para*  
A:Reference number: A41654; MUID:92041655; PMID:1938942  
A:Accession: D41654  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-4 <KRO>  
A:Cross-references: UNIPARC:UPI000017AB26

Query Match 55.6%; Score 5; DB 2; Length 4;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1  
|  
Db 2 R 2

RESULT 6  
I40870  
phospholipase C (EC 3.1.4.3) - *Clostridium perfringens* (fragment)  
C:Species: *Clostridium perfringens*  
C:Date: 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 21-Jul-2000  
C:Accession: I40870  
R;Toyonaga, T.; Matsushita, O.; Katayama, S.; Minami, J.; Okabe, A.  
Microbiol. Immunol. 36, 603-613, 1992  
A:Title: Role of the upstream region containing an intrinsic DNA curvature in the negative  
A:Reference number: I40870; MUID:92396045; PMID:1522810  
A:Accession: I40870  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-4 <RES>  
A:Cross-references: UNIPARC:UPI000011E99B; EMBL:X62825; NID:940622; PIDN:CAA44636.1; PID

C:Genetics:  
A:Gene: plc  
C:Keywords: phosphoric diester hydrolase

Query Match 55.6%; Score 5; DB 2; Length 4;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1  
|  
Db 3 R 3

RESULT 7  
A25844  
autho-RF amide neuropeptide - sea pansy (*Renilla koellikeri*)  
C:Species: *Renilla koellikeri* (Koelliker's sea pansy)  
C:Date: 21-May-1988 #sequence\_revision 30-Sep-1993 #text\_change 11-Jul-1997  
C:Accession: A25844  
R;Grimmelikhuijzen, C.J.P.; Groeger, A.  
FEBS Lett. 211, 105-108, 1987  
A:Title: Isolation of the neuropeptide pGlu-Gly-Arg-Phe-amide from the pennatulid *Renilla*  
A:Reference number: A25844  
A:Accession: A25844  
A:Molecule type: protein  
A:Residues: 1-4 <GRI>  
A:Cross-references: UNIPARC:UPI00001733AF

C:Keywords: amidated carboxyl end; neuropeptide; pyroglutamic acid  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F:4/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 55.6%; Score 5; DB 2; Length 4;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1  
|  
Db 3 R 3

RESULT 8  
I61883

protamine P1 - orangutan (fragment)  
C:Species: *Pongo pygmaeus* (orangutan)  
C:Date: 06-Sep-1996 #sequence\_revision 06-Sep-1996 #text\_change 21-Jul-2000  
C:Accession: I61883  
R;Queralt, R.; Oliva, R.  
Gene 133, 197-204, 1993  
A:Title: Identification of conserved potential regulatory sequences of the protamine-en  
A:Reference number: I37013; MUID:94040810; PMID:8224908  
A:Accession: I61883  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-4 <RES>  
A:Cross-references: UNIPARC:UPI000011E9E7; EMBL:Z12146; NID:938156; PIDN:CAA78130.1; PID

Query Match 55.6%; Score 5; DB 2; Length 4;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1  
|  
Db 3 R 3

RESULT 9  
S47552  
ubiquitin - rat  
C:Species: *Rattus norvegicus* (Norway rat)  
C:Date: 07-May-1995 #sequence\_revision 21-Jul-1995 #text\_change 17-Mar-1999  
C:Accession: S47552  
R;Hubbard, M.J.; Carne, A.  
Biochim. Biophys. Acta 1200, 191-196, 1994  
A:Title: Differential feeding-related regulation of ubiquitin and calbindin(9kDa) in rat  
A:Reference number: S47552; MUID:94304928; PMID:8031840  
A:Accession: S47552  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-4 <HUB>  
A:Cross-references: UNIPARC:UPI000015207C

Query Match 55.6%; Score 5; DB 2; Length 4;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1  
|  
Db 2 R 2

RESULT 10  
I37013  
protamine P1 - *Cercopithecus patas* (fragment)  
C:Species: *Cercopithecus patas*  
C:Date: 19-Mar-1997 #sequence\_revision 07-Nov-1997 #text\_change 21-Jul-2000  
C:Accession: I37013  
R;Queralt, R.; Oliva, R.  
Gene 133, 197-204, 1993  
A:Title: Identification of conserved potential regulatory sequences of the protamine-en  
A:Reference number: I37013; MUID:94040810; PMID:8224908  
A:Accession: I37013  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-4 <RES>  
A:Cross-references: UNIPARC:UPI000011E9E7; EMBL:Z12150; NID:922814; PIDN:CAA78134.1; PID

Query Match 55.6%; Score 5; DB 2; Length 4;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1  
|

Db 3 R 3

RESULT 11

I84439

protamine P1 - savannah baboon (fragment)

C:Species: Papio hamadryas doguera (savannah baboon)

C>Date: 19-Mar-1997 #sequence\_revision 07-Nov-1997 #text\_change 21-Jul-2000

C:Accession: I84439

R:Queralt, R.; Oliva, R.

Gene 133, 197-204, 1993

A:Title: Identification of conserved potential regulatory sequences of the protamine-end

A:Reference number: 137013; MUID:94040810; PMID:8224908

A:Accession: I84439

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-4 <RES>

A:Cross-references: UNIPARC:UPI000011E9E7; EMBL:Z12147; NID:g38134; PIDN:CAA78131.1; PID

Query Match 55.6%; Score 5; DB 2; Length 4;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1

Db 3 R 3

RESULT 12

A35779

neuropeptide Antho-RNamide - sea anemone (Anthopleura elegantissima)

C:Species: Anthopleura elegantissima

C>Date: 04-Dec-1992 #sequence\_revision 04-Dec-1992 #text\_change 09-Jul-2004

C:Accession: A35779

R:Grimmelikhuijzen, C.J.P.; Rinehart, K.L.; Jacob, E.; Graff, D.; Reinscheid, R.K.; Noth

Proc. Natl. Acad. Sci. U.S.A. 87, 5410-5414, 1990

A:Title: Isolation of L-3-phenylacetyl-Leu-Arg-Ans-NH2 (Antho-RNamide), a sea anemone ne

A:Reference number: A35779; MUID:90319122; PMID:1973541

A:Accession: A35779

A:Molecule type: protein

A:Residues: 1-4 <GRI>

A:Cross-references: UNIPROT:P58707; UNIPARC:UPI000012AA36

C:Comment: The L-3-phenylacetyl blocking group probably arises from an amino-terminal ph

C:Keywords: amidated carboxyl end; neuropeptide; phenylacetylation

F:1/Modified site: L-3-phenylacetic acid (Phe) #status experimental

F:4/Modified site: amidated carboxyl end (Asn) #status experimental

Query Match 55.6%; Score 5; DB 2; Length 4;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1

Db 3 R 3

RESULT 13

A60418

FMRFamide - polychaete (Nereis virens)

C:Species: Nereis virens (sandworm)

C>Date: 11-Feb-1993 #sequence\_revision 11-Feb-1993 #text\_change 09-Jul-2004

C:Accession: A60418

R:Krajinak, K.G.; Price, D.A.

Peptides 11, 75-77, 1990

A:Title: Authentic FMRFamide is present in the polychaete Nereis virens.

A:Reference number: A60418; MUID:90259866; PMID:2342992

A:Accession: A60418

A:Molecule type: protein

A:Residues: 1-4 <KRA>

A:Cross-references: UNIPROT:P01162; UNIPARC:UPI000012AAD5

C:Keywords: amidated carboxyl end; neuropeptide

F:4/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 55.6%; Score 5; DB 2; Length 4;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1

Db 3 R 3

RESULT 14

PT0721

T-cell receptor beta chain V-D-J region (140-2K) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 09-Jul-2004

C:Accession: PT0721

R:Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A:Reference number: PT0509; MUID:91277601; PMID:1711558

A:Accession: PT0721

A>Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-4 <FEE>

A:Cross-references: UNIPROT:Q8JZN5; UNIPROT:P54729; UNIPROT:Q8CBY1; UNIPROT:Q8COC0; UNIP

A:Experimental source: newborn thymus, strain BALB/c

C:Keywords: T-cell receptor

Query Match 55.6%; Score 5; DB 2; Length 4;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1

Db 2 R 2

RESULT 15

ECNK

cardioexcitatory neuropeptide FMRFamide - sunray clam

C:Species: Macrocallista nimbosa (sunray clam)

C>Date: 20-Jun-2000 #sequence\_revision 20-Jun-2000 #text\_change 16-Aug-2004

C:Accession: A01426

R:Price, D.A.; Greenberg, M.J.

Science 197, 670-671, 1977

A:Title: Structure of a molluscan cardioexcitatory neuropeptide.

A:Reference number: A01426; MUID:77215956; PMID:877582

A:Accession: A01426

A:Molecule type: protein

A:Residues: 1-4 <PRI>

A:Cross-references: UNIPROT:P01162; UNIPARC:UPI000012AAD5

A:Note: the active peptide was also synthesized

C:Comment: This peptide was purified from pooled extracts of cerebral, pedal, and viscer

C:Keywords: amidated carboxyl end; neuropeptide

F:4/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 55.6%; Score 5; DB 2; Length 4;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1

Db 3 R 3

RESULT 16

H0ROHA

proctolin - American cockroach

C:Species: Periplaneta americana (American cockroach)

C>Date: 29-Jul-1981 #sequence\_revision 29-Jul-1981 #text\_change 09-Jul-2004

C:Accession: A01644

R:Startratt, A.N.; Brown, B.E.

Life Sci. 17, 1253-1256, 1975

Query Match 55.6%; Score 5; DB 2; Length 4;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1

Db 3 R 3

RESULT 14

PT0721

T-cell receptor beta chain V-D-J region (140-2K) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 09-Jul-2004

C:Accession: PT0721

R:Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A:Reference number: PT0509; MUID:91277601; PMID:1711558

A:Accession: PT0721

A>Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-4 <FEE>

A:Cross-references: UNIPROT:Q8JZN5; UNIPROT:P54729; UNIPROT:Q8CBY1; UNIPROT:Q8COC0; UNIP

A:Experimental source: newborn thymus, strain BALB/c

C:Keywords: T-cell receptor

Query Match 55.6%; Score 5; DB 2; Length 4;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1

Db 2 R 2

RESULT 15

ECNK

cardioexcitatory neuropeptide FMRFamide - sunray clam

C:Species: Macrocallista nimbosa (sunray clam)

C>Date: 20-Jun-2000 #sequence\_revision 20-Jun-2000 #text\_change 16-Aug-2004

C:Accession: A01426

R:Price, D.A.; Greenberg, M.J.

Science 197, 670-671, 1977

A:Title: Structure of a molluscan cardioexcitatory neuropeptide.

A:Reference number: A01426; MUID:77215956; PMID:877582

A:Accession: A01426

A:Molecule type: protein

A:Residues: 1-4 <PRI>

A:Cross-references: UNIPROT:P01162; UNIPARC:UPI000012AAD5

A:Note: the active peptide was also synthesized

C:Comment: This peptide was purified from pooled extracts of cerebral, pedal, and viscer

C:Keywords: amidated carboxyl end; neuropeptide

F:4/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 55.6%; Score 5; DB 2; Length 4;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1

Db 3 R 3

RESULT 16

H0ROHA

proctolin - American cockroach

C:Species: Periplaneta americana (American cockroach)

C>Date: 29-Jul-1981 #sequence\_revision 29-Jul-1981 #text\_change 09-Jul-2004

C:Accession: A01644

R:Startratt, A.N.; Brown, B.E.

Life Sci. 17, 1253-1256, 1975

A;Title: Structure of the pentapeptide proctolin, a proposed neurotransmitter in insects  
A;Reference number: A93048; MUID:76074708; PMID:576  
A;Accession: A01644  
A;Molecule type: protein  
A;Residues: 1-5 <STA>  
A;Cross-references: UNIPROT:P01373; UNIPARC:UPI0000132177  
A;Note: the synthetic peptide had the same chromatographic, electrophoretic, and pharmacological properties as the natural peptide.  
R;O'Shea, M.; Adams, M.E.  
Science 213, 567-569, 1981  
A;Title: Pentapeptide (proctolin) associated with an identified neuron.  
A;Reference number: A94260; MUID:81225865; PMID:61113690  
A;Contents: annotation; biological source  
C;Comment: This peptide is found in the lateral white neurons, which occur (in the cockroach) innervate the striated hindgut muscles in insects and stimulate contraction of these muscles.  
C;Superfamily: proctolin  
C;Keywords: neuropeptide

Query Match 55.6%; Score 5; DB 1; Length 5;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1  
|  
Db 1 R 1

RESULT 17  
JN0862  
peptidyl-diesterase A inhibitory peptide C112 - striped bonito  
C;Species: Sarda orientalis (striped bonito)  
C;Date: 10-Mar-1994 #sequence\_revision 10-Mar-1994 #text\_change 07-May-1999  
C;Accession: JN0862  
R;Matsumura, N.; Fujii, M.; Takeda, Y.; Shimizu, T.  
Biosci. Biotechnol. Biochem. 57, 1743-1744, 1993  
A;Title: Isolation and characterization of angiotensin I-converting enzyme inhibitory peptide  
A;Reference number: JN0859; MUID:94080036; PMID:7764272  
A;Accession: JN0862  
A;Molecule type: protein  
A;Residues: 1-5 <MAT>  
A;Cross-references: UNIPARC:UPI00001567C3  
A;Experimental source: intestine  
C;Comment: The amino terminal tripeptide of this protein inhibits angiotensin I-converting enzyme  
C;Superfamily: bradykinin-potentiating peptide  
C;Keywords: angiotensin-converting enzyme inhibitor

Query Match 55.6%; Score 5; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1  
|  
Db 2 R 2

RESULT 18  
I40702  
primase - Citrobacter diversus (fragment)  
C;Species: Citrobacter diversus  
C;Date: 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 16-Aug-1996  
C;Accession: I40702  
R;Versalovic, J.; Lupski, J.R.  
Mol. Microbiol. 8, 343-355, 1993  
A;Title: Conservation and evolution of the rpsU-dnaG-rpoD macromolecular synthesis (MMS) genes  
A;Reference number: I40702; MUID:93302510; PMID:8316085  
A;Accession: I40702  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-5 <RES>  
A;Cross-references: UNIPARC:UPI000011EC67; GB:L01754; NID:g144439  
C;Genetics:  
A;Gene: dnaG

Query Match 55.6%; Score 5; DB 2; Length 5;

Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 R 1  
|  
Db 4 R 4  
RESULT 19  
A44955  
alkaline monooxygenase (FMN-linked) (EC 1.14.14.3) alpha chain - Vibrio harveyi (fragment)  
C;Species: Vibrio harveyi  
C;Date: 03-Jun-1993 #sequence\_revision 03-Jun-1993 #text\_change 26-May-2000  
C;Accession: A44955  
R;Paquette, O.; Tu, S.C.  
Photochem. Photobiol. 50, 817-825, 1989  
A;Title: Chemical modification and characterization of the alpha cysteine 106 at the V106 site of Vibrio harveyi  
A;Reference number: A44955; MUID:90175700; PMID:2626493  
A;Accession: A44955  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-5 <PAQ>  
A;Cross-references: UNIPARC:UPI000017AAD8  
C;Keywords: FMN; luminescence; monooxygenase; oxidoreductase

Query Match 55.6%; Score 5; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1  
|  
Db 5 R 5

RESULT 20  
I39964  
ribosomal protein S4 - Bacillus circulans (fragment)  
C;Species: Bacillus circulans  
C;Date: 19-Jul-1996 #sequence\_revision 19-Jul-1996 #text\_change 19-Jul-1996  
C;Accession: I39964  
R;Grundy, F.J.; Henkin, T.M.  
J. Bacteriol. 174, 6763-6770, 1992  
A;Title: Characterization of the Bacillus subtilis rpsD regulatory target site.  
A;Reference number: I39963; MUID:93015735; PMID:1400226  
A;Accession: I39964  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-5 <RES>  
A;Cross-references: UNIPARC:UPI000011EC5E; GB:M99041; NID:g143471  
C;Genetics:  
A;Gene: rpsD

Query Match 55.6%; Score 5; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1  
|  
Db 3 R 3

RESULT 21  
I39966  
ribosomal protein S4 - Bacillus licheniformis (fragment)  
C;Species: Bacillus licheniformis  
C;Date: 19-Jul-1996 #sequence\_revision 19-Jul-1996 #text\_change 19-Jul-1996  
C;Accession: I39966  
R;Grundy, F.J.; Henkin, T.M.  
J. Bacteriol. 174, 6763-6770, 1992  
A;Title: Characterization of the Bacillus subtilis rpsD regulatory target site.  
A;Reference number: I39963; MUID:93015735; PMID:1400226  
A;Accession: I39966  
A;Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-5 <RES>

A:Cross-references: UNIPARC:UPI000011EC5E; GB:M99043; NID:g143475

C:Genetics:

A:Gene: rpsD

Query Match 55.6%; Score 5; DB 2; Length 5;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1

Db 3 R 3

RESULT 22

I39965

ribosomal protein S4 - Bacillus megaterium (fragment)

C:Species: Bacillus megaterium

C:Date: 19-Jul-1996 #sequence\_revision 19-Jul-1996 #text\_change 19-Jul-1996

C:Accession: I39965

R:Grundy, F.J.; Henkin, T.M.

J. Bacteriol. 174, 6763-6770, 1992

A:Title: Characterization of the Bacillus subtilis rpsD regulatory target site.

A:Reference number: I39963; MUID:93015735; PMID:1400226

A:Accession: I39965

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-5 <RES>

A:Cross-references: UNIPARC:UPI000011EC5E; GB:M99042; NID:g143473

C:Genetics:

A:Gene: rpsD

Query Match 55.6%; Score 5; DB 2; Length 5;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1

Db 3 R 3

RESULT 23

D60274

major protein antigen MPT46 - Mycobacterium tuberculosis (fragment)

C:Species: Mycobacterium tuberculosis

C:Date: 11-Dec-1992 #sequence\_revision 11-Dec-1992 #text\_change 30-Sep-1993

C:Accession: D60274

R:Nagai, S.; Wiker, H.G.; Harboe, M.; Kinomoto, M.

Infect. Immun. 59, 372-382, 1991

A:Title: Isolation and partial characterization of major protein antigens in the culture

A:Reference number: A60274; MUID:91099989; PMID:1898899

A:Accession: D60274

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-5 <NAG>

A:Cross-references: UNIPARC:UPI0000150742

Query Match 55.6%; Score 5; DB 2; Length 5;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1

Db 1 R 1

RESULT 24

B22565

R-phycoerythrin alpha-2 chain - red alga (Gastrocloonium coulteri) (fragment)

C:Species: Gastrocloonium coulteri

C:Date: 07-Mar-1988 #sequence\_revision 07-Mar-1988 #text\_change 23-Mar-1993

C:Accession: B22565

R:Klotz, A.V.; Glazer, A.N.

J. Biol. Chem. 260, 4856-4863, 1985

A:Title: Characterization of the bilin attachment sites in R-phycoerythrin.

A:Reference number: A22565; MUID:85182601; PMID:3886644

A:Accession: B22565

A:Molecule type: protein

A:Residues: 1-5 <KLO>

A:Cross-references: UNIPARC:UPI000017AEC3

Query Match 55.6%; Score 5; DB 2; Length 5;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1

Db 5 R 5

RESULT 25

F22565

R-phycoerythrin gamma-A chain - red alga (Gastrocloonium coulteri) (fragment)

C:Species: Gastrocloonium coulteri

C:Date: 07-Mar-1988 #sequence\_revision 07-Mar-1988 #text\_change 23-Mar-1993

C:Accession: F22565

R:Klotz, A.V.; Glazer, A.N.

J. Biol. Chem. 260, 4856-4863, 1985

A:Title: Characterization of the bilin attachment sites in R-phycoerythrin.

A:Reference number: A22565; MUID:85182601; PMID:3886644

A:Accession: F22565

A:Molecule type: protein

A:Residues: 1-5 <KLO>

A:Cross-references: UNIPARC:UPI000017AEC4

Query Match 55.6%; Score 5; DB 2; Length 5;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1

Db 5 R 5

RESULT 26

T14910

hypothetical protein - parsley

C:Species: Petroselinum crispum (parsley)

C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 21-Jul-2000

C:Accession: T14910

R:Kircher, S.; Ledger, S.; Havaehi, H.; Weisshaar, B.; Schafer, E.; Frohnmeyer, H.

Mol. Gen. Genet. 257, 595-605, 1998

A:Title: CPRP4a, a novel plant BZIP protein of the CPRF family: comparative analysis of

A:Reference number: Z18261; MUID:98265918; PMID:9604882

A:Accession: T14910

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-5 <KIR>

A:Cross-references: UNIPARC:UPI000011E9D7; EMBL:Y10810; NID:g3336904; PIDN:CAA71769.1;

A:Experimental source: ssp. Hamburger Schnitt

Query Match 55.6%; Score 5; DB 2; Length 5;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1

Db 4 R 4

RESULT 27

A60803

neuropeptide - sea anemone (Anthopleura elegantissima)

C:Species: Anthopleura elegantissima

C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 07-May-1999

C:Accession: B22565

C:Accession: A60803  
 R:Graff, D.; Grimelikhuijzen, C.J.P.  
 Brain Res. 442, 354-358, 1988

A:Title: Isolation of <Glu-Ser-Lu-Arg-Tyr-NH-2, a novel neuropeptide from sea anemones.  
 A:Reference number: A60803; MUID:88222764; PMID:2897223

C:Accession: A60803  
 A:Molecule type: protein  
 A:Residues: 1-5 <GR>

A:Cross-references: UNIPARC:UPI000017B683  
 C:Keywords: amidated carboxyl end; neuropeptide; pyroglutamic acid  
 F1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
 F5/Modified site: amidated carboxyl end (trp) #status experimental

Query Match 55.6%; Score 5; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1  
 |  
 Db 4 R 4

## RESULT 28

S53595

hypothetical protein (upstream of transcription factor, CCAAT-binding) - chicken

C:Species: Gallus gallus (chicken)

C:Date: 15-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 07-May-1999

C:Accession: S53595

R:Calikhoven, C.F.; Bouwman, P.R.J.; Snippe, L.; Ab, G.

Nucleic Acids Res. 22, 5540-5547, 1994

A:Title: Translation start site multiplicity of the CCAAT/enhancer binding protein alpha

A:Reference number: S53595; MUID:95140613; PMID:7838705

C:Accession: S53595

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-5 <CAL>

A:Cross-references: UNIPARC:UPI000017C010; EMBL:X66844

Query Match 55.6%; Score 5; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1  
 |  
 Db 4 R 4

## RESULT 29

PT0295

Ig heavy chain CRD3 region (clone 5-91) - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996

C:Accession: PT0295

R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.

J. Exp. Med. 173, 395-407, 1991

A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity and

A:Reference number: PT0222; MUID:91108337; PMID:1899102

C:Accession: PT0295

A:Molecule type: DNA

A:Residues: 1-5 <YAM>

A:Cross-references: UNIPARC:UPI000017C20B

A:Experimental source: B lymphocyte

C:Keywords: heterotetramer; immunoglobulin

Query Match 55.6%; Score 5; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1  
 |  
 Db 4 R 4

## RESULT 30

S62883

seminal plasma protein II - pig (fragment)

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change 17-Mar-1999

C:Accession: S62883

R:Romero, A.; Varela, P.F.; Sanz, L.; Toepfer-Petersen, E.; Calvete, J.J.

FEBS Lett. 382, 15-17, 1996

A:Title: Crystallization and preliminary X-ray diffraction analysis of boar seminal pla

A:Reference number: S62882; MUID:96196555; PMID:8612739

C:Accession: S62883

A:Molecule type: protein

A:Residues: 1-5 <ROM>

A:Cross-references: UNIPARC:UPI000017C475

C:Complex: heterodimer; seminal plasma protein I and seminal plasma protein II

C:Keywords: glycoprotein; heterodimer; semen

Query Match 55.6%; Score 5; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1  
 |  
 Db 2 R 2

## RESULT 31

PT0513

T-cell receptor beta chain V-D-J region (100-4AL) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997

C:Accession: PT0513; PT0606

R:Peeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A:Reference number: PT0509; MUID:91277601; PMID:1711558

C:Accession: PT0513

A>Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-5 <FEE>

A:Cross-references: UNIPARC:UPI000017C78C

A:Experimental source: adult thymus, strain BALB/c, clone 100-4AL

C:Accession: PT0606

A>Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-5 <PE2>

A:Cross-references: UNIPARC:UPI000017C78C

A:Experimental source: newborn thymus, strain BALB/c, clone 120-1S

C:Keywords: T-cell receptor

Query Match 55.6%; Score 5; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1  
 |  
 Db 5 R 5

## RESULT 32

PT0525

T-cell receptor beta chain V-D-J region (100-4J) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997

C:Accession: PT0525

R:Peeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A:Reference number: PT0509; MUID:91277601; PMID:1711558

C:Accession: PT0525

A>Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-5 <FEE>

A;Cross-references: UNIPARC:UPI000017C7C4  
A;Experimental source: adult thymus, strain BALB/c  
C;Keywords: T-cell receptor

Query Match 55.6%; Score 5; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1  
|  
Db 3 R 3

## RESULT 33

PT0597 T-cell receptor beta chain V-D-J region (111-1B) - mouse (fragment)

C;Species: Mus musculus (house mouse)  
C;Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
C;Accession: PT0597

R;Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A;Reference number: PT0509; MUID:91277601; PMID:1711558

A;Accession: PT0597

A;Status: translation not shown

A;Molecule type: mRNA

A;Residues: 1-5 <FEE>

A;Cross-references: UNIPARC:UPI000017C7CF

A;Experimental source: newborn thymus, strain BALB/c

C;Keywords: T-cell receptor

Query Match 55.6%; Score 5; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1  
|  
Db 5 R 5

## RESULT 34

PT0672 T-cell receptor beta chain V-D-J region (121-1B) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997

C;Accession: PT0646; PT0672

R;Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A;Reference number: PT0509; MUID:91277601; PMID:1711558

A;Accession: PT0646

A;Status: translation not shown

A;Molecule type: mRNA

A;Residues: 1-5 <FEE>

A;Cross-references: UNIPARC:UPI000017C7E9

A;Experimental source: day 4 postnatal thymus, strain BALB/c, clone 121-1B

A;Accession: PT0672

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-5 <FE2>

A;Cross-references: UNIPARC:UPI000017C7E9

A;Experimental source: day 18 fetal thymus, strain BALB/c, clone 140-1B

C;Keywords: T-cell receptor

Query Match 55.6%; Score 5; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1  
|  
Db 5 R 5

## RESULT 35

PT0553 T-cell receptor beta chain V-D-J region (126-1C) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997

C;Accession: PT0553

R;Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A;Reference number: PT0509; MUID:91277601; PMID:1711558

A;Accession: PT0553

A;Status: translation not shown

A;Molecule type: mRNA

A;Residues: 1-5 <FEE>

A;Cross-references: UNIPARC:UPI000017C810

A;Experimental source: day 18 fetal thymus, strain BALB/c

C;Keywords: T-cell receptor

Query Match 55.6%; Score 5; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1  
|  
Db 4 R 4

## RESULT 36

PT0695 T-cell receptor beta chain V-D-J region (135-1D) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997

C;Accession: PT0695

R;Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A;Reference number: PT0509; MUID:91277601; PMID:1711558

A;Accession: PT0695

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-5 <PEE>

A;Cross-references: UNIPARC:UPI000017C81B

A;Experimental source: newborn thymus, strain BALB/c

C;Keywords: T-cell receptor

Query Match 55.6%; Score 5; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1  
|  
Db 4 R 4

## RESULT 37

PT0577 T-cell receptor beta chain V-D-J region (141-1BC) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997

C;Accession: PT0577; PT0574

R;Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A;Reference number: PT0509; MUID:91277601; PMID:1711558

A;Accession: PT0577

A;Status: translation not shown

A;Molecule type: mRNA

A;Residues: 1-5 <FEE>

A;Cross-references: UNIPARC:UPI000017C830

A;Experimental source: day 19 fetal thymus, strain BALB/c, clone 141-1BC

A;Accession: PT0574

A;Status: translation not shown

A;Molecule type: mRNA



A;Residues: 1-5 <PE2>  
A;Cross-references: UNIPARC:UPI000017C830  
A;Experimental source: day 19 fetal thymus, strain BALB/c, clone 141-1Q  
C;Keywords: T-cell receptor

Query Match 55.6%; Score 5; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1  
|  
Db 4 R 4

RESULT 38  
PT0572  
T-cell receptor beta chain V-D-J region (141-1CO) - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
C;Accession: PT0572  
R;Feeney, A.J.  
J. Exp. Med. 174, 115-124, 1991  
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
A;Reference number: PT0509; MUID:91277601; PMID:1711558  
A;Accession: PT0572  
A;Status: translation not shown  
A;Molecule type: mRNA  
A;Residues: 1-5 <PEE>  
A;Cross-references: UNIPARC:UPI000017C838  
A;Experimental source: day 19 fetal thymus, strain BALB/c  
C;Keywords: T-cell receptor

Query Match 55.6%; Score 5; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1  
|  
Db 5 R 5

RESULT 39  
PT0700  
T-cell receptor beta chain V-D-J region (161-2A) - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
C;Accession: PT0700  
R;Feeney, A.J.  
J. Exp. Med. 174, 115-124, 1991  
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
A;Reference number: PT0509; MUID:91277601; PMID:1711558  
A;Accession: PT0700  
A;Status: translation not shown  
A;Molecule type: DNA  
A;Residues: 1-5 <PEE>  
A;Cross-references: UNIPARC:UPI000017C841  
A;Experimental source: newborn thymus, strain BALB/c  
C;Keywords: T-cell receptor

Query Match 55.6%; Score 5; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1  
|  
Db 4 R 4

RESULT 40  
A60411  
proctolin - Atlantic horseshoe crab  
C;Species: Limulus polyphemus (Atlantic horseshoe crab)  
C;Date: 03-Feb-1993 #sequence\_revision 03-Feb-1993 #text\_change 09-Jul-2004

C;Accession: A60411  
R;Grome, J.R.; Tillinghast, E.K.; Townley, M.A.; Vetrovs, A.; Watson III, W.H.; Hunt, I.  
Peptides 11, 205-211, 1990  
A;Title: Identification of proctolin in the central nervous system of the horseshoe crab  
A;Reference number: A60411; MUID:90287800; PMID:2356151  
A;Accession: A60411  
A;Molecule type: protein  
A;Residues: 1-5 <GRO>  
A;Cross-references: UNIPROT:P01373; UNIPARC:UPI0000132177  
C;Comment: This neuropeptide stimulates cardiac output and hindgut motility in the horseshoe crab  
C;Keywords: neuropeptide

Query Match 55.6%; Score 5; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1  
|  
Db 1 R 1

RESULT 41  
PT0608  
T-cell receptor beta chain V-D-J region (120-2CP) - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 09-Jul-2004  
C;Accession: PT0608  
R;Feeney, A.J.  
J. Exp. Med. 174, 115-124, 1991  
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
A;Reference number: PT0509; MUID:91277601; PMID:1711558  
A;Accession: PT0608  
A;Status: translation not shown  
A;Molecule type: mRNA  
A;Residues: 1-5 <PEE>  
A;Cross-references: UNIPROT:O18345; UNIPARC:UPI000017C7E0  
A;Experimental source: newborn thymus, strain BALB/c  
C;Keywords: T-cell receptor

Query Match 55.6%; Score 5; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1  
|  
Db 4 R 4

RESULT 42  
PT0565  
T-cell receptor beta chain V-D-J region (141-1CF) - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 09-Jul-2004  
C;Accession: PT0565  
R;Feeney, A.J.  
J. Exp. Med. 174, 115-124, 1991  
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
A;Reference number: PT0509; MUID:91277601; PMID:1711558  
A;Accession: PT0565  
A;Status: translation not shown  
A;Molecule type: mRNA  
A;Residues: 1-5 <PEE>  
A;Cross-references: UNIPROT:Q8BZQ7; UNIPARC:UPI000017C835  
A;Experimental source: day 19 fetal thymus, strain BALB/c  
C;Keywords: T-cell receptor

Query Match 55.6%; Score 5; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1  
|  
Db 4 R 4

```
RESULT 43
A35890
RNA-directed DNA polymerase (EC 2.7.7.49) 66K chain - human immunodeficiency virus type
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 09-Nov-1990 #sequence_revision 09-Nov-1990 #text_change 31-Dec-1993
C:Accession: A35890
R:Batchurst, I.C.; Moen, L.K.; Lujan, M.A.; Gibson, H.L.; Feucht, P.H.; Pichuanes, S.; C
Biochem. Biophys. Res. Commun. 171, 589-595, 1990
A:Title: Characterization of the human immunodeficiency virus type-1 reverse transcripta
A:Reference number: A35890; MUID:90386627; PMID:1698361
A:Accession: A35890
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-6 <BAT>
A:Cross-references: UNIPARC:UPI000017A88C
C:Keywords: nucleotidyltransferase

Query Match 55.6%; Score 5; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1
DB 3 R 3

RESULT 44
A37765
hypothetical protein (csma 5' region) - Chloroflexus aurantiacus (fragment)
C:Species: Chloroflexus aurantiacus
C:Date: 31-May-1991 #sequence_revision 31-May-1991 #text_change 30-Sep-1993
C:Accession: A37765
R:Theroux, S.J.; Redlinger, T.E.; Fuller, R.C.; Robinson, S.J.
J. Bacteriol. 172, 4497-4504, 1990
A:Title: Gene encoding the 5.7-kilodalton chlorosome protein of Chloroflexus aurantiacus
A:Reference number: A37765; MUID:90330558; PMID:2376566
A:Accession: A37765
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-6 <THE>
A:Cross-references: UNIPARC:UPI000017ABN1; GB:M33964

Query Match 55.6%; Score 5; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1
DB 6 R 6

RESULT 45
C22565
R-phycoerythrin beta-1 chain - red alga (Gastroclonium coulteri) (fragment)
C:Species: Gastroclonium coulteri
C:Date: 07-Mar-1988 #sequence_revision 07-Mar-1988 #text_change 23-Mar-1993
C:Accession: C22565
R:Klotz, A.V.; Glazer, A.N.
J. Biol. Chem. 260, 4856-4863, 1985
A:Title: Characterization of the bilin attachment sites in R-phycoerythrin.
A:Reference number: A22565; MUID:85182601; PMID:3886644
A:Accession: C22565
A:Molecule type: protein
A:Residues: 1-6 <KLO>
A:Cross-references: UNIPARC:UPI0000171E4A

Query Match 55.6%; Score 5; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1
```

DB 6 R 6

```
RESULT 46
PQ0008
angiotensin-converting enzyme inhibitor (FLP-1) - common fig
N:Alternate names: ficus latex peptide 1
C:Species: Ficus carica (common fig)
C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 08-Dec-1995
C:Accession: PQ0008
R:Maruyama, S.; Miyoshi, S.; Tanaka, H.
Agric. Biol. Chem. 53, 2763-2767, 1989
A:Title: Angiotensin I-converting enzyme inhibitors derived from Ficus carica.
A:Reference number: PQ0008
A:Accession: PQ0008
A:Molecule type: protein
A:Residues: 1-6 <MAR>
A:Cross-references: UNIPARC:UPI000015655D
A:Experimental source: latex
C:Keywords: angiotensin-converting enzyme inhibitor
```

```
Query Match 55.6%; Score 5; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1
DB 6 R 6
```

```
RESULT 47
A60494
antineoplastic glycoprotein - sea hare (Dolabella auricularia) (fragment)
N:Alternate names: dolabellarin C
C:Species: Dolabella auricularia
C:Date: 19-Mar-1993 #sequence_revision 19-Mar-1993 #text_change 18-Jun-1993
C:Accession: A60494
R:Kisugi, J.; Kamiya, H.; Yamazaki, M.
Dev. Comp. Immunol. 13, 3-8, 1989
A:Title: Purification of dolabellarin-C an antineoplastic glycoprotein in the body fluid
A:Reference number: A60494; MUID:89357188; PMID:2767307
A:Accession: A60494
A:Molecule type: protein
A:Residues: 1-6 <KIS>
A:Cross-references: UNIPARC:UPI000017BD9D
C:Keywords: cytotoxicity; glycoprotein; trimer
```

```
Query Match 55.6%; Score 5; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1
DB 5 R 5
```

```
RESULT 48
I51434
H4 histone - African clawed frog (fragment)
C:Species: Xenopus laevis (African clawed frog)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C:Accession: I51434
R:Woodland, H.R.; Warrington, J.R.; Ballantine, J.E.M.; Turner, P.C.
Nucleic Acids Res. 12, 4939-4958, 1984
A:Title: Are there major developmentally regulated H4 gene classes in Xenopus?
A:Reference number: I51391; MUID:84247348; PMID:6330691
A:Accession: I51434
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-6 <WOO>
A:Cross-references: UNIPARC:UPI000011E896; GB:K02304; NID:G214227; PIDN:AAA49738.1; PID
```

Query Match 55.6%; Score 5; DB 2; Length 6;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1  
|  
Db 4 R 4

RESULT 49  
I37027  
protamine P1 - gorilla (fragment)  
C:Species: Gorilla gorilla (gorilla)  
C:Date: 04-Oct-1996 #sequence\_revision 04-Oct-1996 #text\_change 21-Jul-2000  
C:Accession: I37027  
R:Queralt, R.; Oliva, R.  
Gene 133, 197-204, 1993  
A:Title: Identification of conserved potential regulatory sequences of the protamine-end  
A:Reference number: I37013; MUID:94040810; PMID:8224908  
A:Accession: I37027  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-6 <RES>  
A:Cross-references: UNIPARC:UPI000011E9E8; EMBL:Z12145; NID:g22910; PIDN:CAA78129.1; PID

Query Match 55.6%; Score 5; DB 2; Length 6;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1  
|  
Db 3 R 3

RESULT 50  
A11490  
pyruvate kinase (EC 2.7.1.40) - pig (fragment)  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 05-Jun-1987 #sequence\_revision 05-Jun-1987 #text\_change 03-Mar-1995  
C:Accession: A11490  
R:Hjelmquist, G.; Andersson, J.; Edlund, B.; Engstrom, L.  
Biochem. Biophys. Res. Commun. 61, 559-563, 1974  
A:Title: Amino acid sequence of a (32-P)phosphopeptide from pig liver pyruvate kinase ph  
A:Reference number: A11490; MUID:75127438; PMID:4375989  
A:Accession: A11490  
A:Molecule type: protein  
A:Residues: 1-6 <HJE>  
A:Cross-references: UNIPARC:UPI000017C474  
A:Experimental source: liver  
C:Keywords: glycolysis; phosphotransferase

Query Match 55.6%; Score 5; DB 2; Length 6;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1  
|  
Db 2 R 2

Search completed: January 25, 2006, 18:41:27  
Job time : 18.5 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:33:23 ; Search time 63 Seconds  
(without alignments)  
55.994 Million cell updates/sec

Title: US-10-771-242-293

Perfect score: 9

Sequence: 1 RXXX 5

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database : UniProt 05.80.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	ID	Description
1	5	55.6	4 1 FAR3_HIRME	P42562 hirudo medi
2	5	55.6	4 1 FAR4_HIRME	P42563 hirudo medi
3	5	55.6	4 1 FLRF_HELTI	P69138 helisoma tr
4	5	55.6	4 1 FLRF_HIRME	P69137 hirudo medi
5	5	55.6	4 1 FLRF_ATEL	P58707 anthopleura
6	5	55.6	4 1 FMRP_HELTI	P69148 helisoma tr
7	5	55.6	4 1 FMRP_HIRME	P69147 hirudo medi
8	5	55.6	4 1 FMRP_MACNI	P69145 macrocallis
9	5	55.6	4 1 FMRP_NERVI	P69146 nereis, vire
10	5	55.6	4 1 TUFT_HUMAN	P58706 anthopleura
11	5	55.6	4 1 TUFT_HUMAN	P01858 homo sapien
12	5	55.6	5 1 FARP_ARTTR	P41853 artiopesthi
13	5	55.6	5 1 FARP_CHICK	P83308 gallus gall
14	5	55.6	5 1 PRCT_CARMA	P67857 carcinus ma
15	5	55.6	5 1 PRCT_LIMPO	P67858 limulus pol
16	5	55.6	5 1 PRCT_PERAM	P67859 periplaneta
17	5	55.6	5 1 UF01_MOUSE	P38639 mus musculu
18	5	55.6	5 1 ACPH_RABIT	P25154 oryctolagus
19	5	55.6	6 1 FARP_MONEX	P41966 moniezia ex
20	5	55.6	6 1 PYF1_PENMO	P84005 penaeus mon
21	5	55.6	6 2 P82181_SPIOL	P82181 spinacia ol
22	5	55.6	6 2 P82182_SPIOL	P82182 spinacia ol
23	5	55.6	6 2 P82541_SPIOL	P82541 spinacia ol
24	5	55.6	7 1 CAP6_CANAL	P83784 candida alb
25	5	55.6	7 1 CARP_MYTED	P10420 mytilus edu
26	5	55.6	7 1 CHOK_ALCSP	P16101 alcaligenes
27	5	55.6	7 1 FAF1_ASCSU	P31889 ascaris suu
28	5	55.6	7 1 FAF2_PANRE	P67879 ascaris suu
29	5	55.6	7 1 FAF2_PANRE	P67880 panagrellus
30	5	55.6	7 1 FARI_HELTI	P41871 helisoma tr
31	5	55.6	7 1 FARI_MACRS	P83274 macrobrachi

32	5	55.6	7 1 FAR1_PROCL	P38499 procambarus
33	5	55.6	7 1 FAR2_PROCL	P38498 procambarus
34	5	55.6	7 1 FAR3_HABCO	P81298 haemochus
35	5	55.6	7 1 FAR4_PANRE	P41874 panagrellus
36	5	55.6	7 1 FAR4_PANRE	P41875 panagrellus
37	5	55.6	7 1 FAR5_HIRME	P42564 hirudo medi
38	5	55.6	7 1 FAR6_CALVO	P41866 calliphora
39	5	55.6	7 1 HCYB_CONCC	P84620 concholepas
40	5	55.6	7 1 IPYR_CANAL	P83777 candida alb
41	5	55.6	7 1 LANC_CARUI	P36960 carnobacter
42	5	55.6	7 1 TV51_LITRU	P82065 litoria rub
43	5	55.6	7 1 UF04_MOUSE	P38642 mus musculu
44	5	55.6	7 1 UH11_RAT	P56576 rattus norv
45	5	55.6	7 1 WWAI_ACHFU	P35919 achatina fu
46	5	55.6	7 2 Q95945_YEAST	Q95945 saccharomyc
47	5	55.6	7 2 Q15903_HUMAN	Q15903 homo sapien
48	5	55.6	7 2 Q8NH7_HUMAN	Q8NH7 homo sapien
49	5	55.6	7 2 O8TAQ4_HUMAN	Q8TAQ4 homo sapien
50	5	55.6	7 2 O8866_SPIOL	O8866 spinacia ol
51	5	55.6	7 2 P92210_AGRCR	P92210 agropyron c
52	5	55.6	7 2 P92214_9FOAL	P92214 amblyopyrum
53	5	55.6	7 2 P92218_9FOAL	P92218 australopyr
54	5	55.6	7 2 P92221_BROIN	P92221 bromus iner
55	5	55.6	7 2 P92226_CRIDE	P92226 crithopsis
56	5	55.6	7 2 P92372_9FOAL	P92372 haynaldia v
57	5	55.6	7 2 P92381_9FOAL	P92381 hordeum bra
58	5	55.6	7 2 P92385_HORMA	P92385 hordeum mar
59	5	55.6	7 2 P92387_9FOAL	P92387 henrardia p
60	5	55.6	7 2 P92390_HETPI	P92390 heteranthel
61	5	55.6	7 2 P92393_HORVU	P92393 hordeum vul
62	5	55.6	7 2 P92403_LOPEL	P92403 lophopyrum
63	5	55.6	7 2 P92421_PSAFR	P92421 psathyrosta
64	5	55.6	7 2 P92425_PSEPI	P92425 pseudoroegn
65	5	55.6	7 2 P92427_9FOAL	P92427 peridictyon
66	5	55.6	7 2 P92430_AEGTA	P92430 aegilops ta
67	5	55.6	7 2 P92440_THIBE	P92440 thinopyrum
68	5	55.6	7 2 P92442_TAECM	P92442 taeniatheru
69	5	55.6	7 2 P93233_LYCES	P93233 lycopersico
70	5	55.6	7 2 P84495_CUCMA	P84495 cucurbita m
71	5	55.6	7 2 O34028_9SPHN	O34028 sphingomona
72	5	55.6	7 2 O50556_ACTAC	O50556 actinobacil
73	5	55.6	7 2 O47477_ECOLI	O47477 escherichia
74	5	55.6	7 2 O07354_SYNP8	O07354 synechococc
75	5	55.6	7 2 Q47505_ECOLI	Q47505 escherichia
76	5	55.6	7 2 Q63480_RAT	Q63480 rattus norv
77	5	55.6	7 2 Q66113_9COMO	Q66113 cherry leaf
78	5	55.6	7 2 Q67113_9NFA	Q67113 influenza a
79	5	55.6	7 2 Q9Y1Q9_ADE04	Q9Y1Q9 human adeno
80	5	55.6	7 2 Q9Y1R0_ADE07	Q9Y1R0 human adeno
81	5	55.6	7 2 Q9YVE3_ADE07	Q9YVE3 carcinus ma
82	5	55.6	8 1 ACT_CARMA	P82154 cydia pomon
83	5	55.6	8 1 ALL3_CVDPO	P82154 cydia pomon
84	5	55.6	8 1 ALL4_CALVO	P41840 calliphora
85	5	55.6	8 1 ALL4_CVDPO	P82155 cydia pomon
86	5	55.6	8 1 ALL5_CVDPO	P82156 cydia pomon
87	5	55.6	8 1 ANGT2_BOTJA	Q10582 bothrops ja
88	5	55.6	8 1 B44K_PORGI	P81886 porphyromon
89	5	55.6	8 1 C125_CVPDO	P83661 cyphononyx
90	5	55.6	8 1 COX6B_RAT	P80430 rattus norv
91	5	55.6	8 1 FAR1_PANRE	P41872 panagrellus
92	5	55.6	8 1 FAR1_PENMO	P83316 penaeus mon
93	5	55.6	8 1 FAR2_MACRS	P83275 macrobrachi
94	5	55.6	8 1 FAR3_HOMAM	P41486 homarus ame
95	5	55.6	8 1 FAR4_HOMAM	P41487 homarus ame
96	5	55.6	8 1 FAR4_MACRS	P83277 macrobrachi
97	5	55.6	8 1 FAR7_ASCSU	P43171 ascaris suu
98	5	55.6	8 1 FAR8_CALVO	P41863 calliphora
99	5	55.6	8 1 KIN11_PERAM	P82685 periplaneta
100	5	55.6	8 1 LMT2_LOEMI	P23296 locusta mig
101	5	55.6	8 1 LPK_LEUMA	P13049 leucophaea
102	5	55.6	8 1 LPM5_STAEP	P23211 staphylococ
103	5	55.6	8 1 NPMB_BOVIN	P15507 bos taurus
104	5	55.6	8 1 NS3_MYCTU	P81152 mycobacteri

105	5	55.6	8	1	PPK2_PERAM	P82692	periplaneta	178	5	55.6	8	2	Q8H9J9_BPM1	Q8h9j9 bacterioph
106	5	55.6	8	1	PPK3_PERAM	P82618	periplaneta	179	5	55.6	8	2	Q8H9K4_BPK3	Q8h9k4 bacterioph
107	5	55.6	8	1	UC7_MYCIT	P33564	mycobacteri	180	5	55.6	8	2	O19956_GOSAR	O19956 gossypium a
108	5	55.6	8	1	UC26_MAIZE	P80632	zea mays (m	181	5	55.6	8	2	O19957_GOSHI	O19957 gossypium b
109	5	55.6	8	1	UP09_RAT	P56575	rattus norv	182	5	55.6	8	2	O19958_GOSBA	O19958 gossypium t
110	5	55.6	8	1	UPA1_HUMAN	P30087	homo sapien	183	5	55.6	8	2	O19959_GOSBO	O19959 gossypium b
111	5	55.6	8	2	Q7M4U4_ASPIFI	O7M4U4	aspergillus	184	5	55.6	8	2	O19960_GOSMU	O19960 gossypium m
112	5	55.6	8	2	Q9T2W0_YEAST	Q9T2W0	saccharomyc	185	5	55.6	8	2	O19961_GOSDA	O19961 gossypium d
113	5	55.6	8	2	Q15888_HUMAN	Q15888	homo sapien	186	5	55.6	8	2	Q36898_9SOLA	Q36898 nicotiana p
114	5	55.6	8	2	Q15898_HUMAN	Q15898	homo sapien	187	5	55.6	8	2	Q5D4X1_9MYRT	Q5d4x1 physocallym
115	5	55.6	8	2	Q15900_HUMAN	Q15900	homo sapien	188	5	55.6	8	2	Q5D4X2_9MYRT	Q5d4x2 pephria comp
116	5	55.6	8	2	Q16468_HUMAN	Q16468	homo sapien	189	5	55.6	8	2	Q5D4X4_9MYRT	Q5d4x4 koehneria m
117	5	55.6	8	2	Q69YH8_HUMAN	Q69YH8	homo sapien	190	5	55.6	8	2	Q5D4X5_9MYRT	Q5d4x5 nesaea aspe
118	5	55.6	8	2	Q710R9_HUMAN	Q710R9	homo sapien	191	5	55.6	8	2	Q5D4Y0_9MYRT	Q5d4y0 lafoensia a
119	5	55.6	8	2	Q724N9_HUMAN	Q724N9	homo sapien	192	5	55.6	8	2	Q5D4Y1_9MYRT	Q5d4y1 ginoria ame
120	5	55.6	8	2	Q81VK3_HUMAN	Q81VK3	homo sapien	193	5	55.6	8	2	Q5D4Y5_9MYRT	Q5d4y5 capuroniam
121	5	55.6	8	2	Q8YF70_HUMAN	Q8YF70	homo sapien	194	5	55.6	8	2	Q5IA47_9MAGN	Q5ia47 laurus nobi
122	5	55.6	8	2	Q9Y4J4_HUMAN	Q9Y4J4	homo sapien	195	5	55.6	8	2	Q5Y9D8_9BRYO	Q5y9d8 warnstorfia
123	5	55.6	8	2	Q539B0_HUMAN	Q539B0	homo sapien	196	5	55.6	8	2	Q659Q3_9CARY	Q659q3 silene oste
124	5	55.6	8	2	Q9UMH9_HUMAN	Q9UMH9	homo sapien	197	5	55.6	8	2	Q659Q5_9CARY	Q659q5 silene invo
125	5	55.6	8	2	Q75MD1_HUMAN	Q75MD1	homo sapien	198	5	55.6	8	2	Q6H0C6_9TRAC	Q6h0c6 isoetes fla
126	5	55.6	8	2	Q4EW64_HUMAN	Q4EW64	homo sapien	199	5	55.6	8	2	Q6YL78_SCIVE	Q6yl78 sciadopitys
127	5	55.6	8	2	O15896_BABBO	O15896	babesia bov	200	5	55.6	8	2	Q6Z202_9CARY	Q6z202 lychnis cor
128	5	55.6	8	2	O15899_BABOV	O15899	babesia ovi	201	5	55.6	8	2	Q70Y68_9LAMI	Q70y68 prostanther
129	5	55.6	8	2	Q70MX3_TRYBR	Q70MX3	trypanosoma	202	5	55.6	8	2	Q70Y69_9LAMI	Q70y69 syncolostem
130	5	55.6	8	2	Q7M3N2_MANSE	Q7M3N2	manduca sex	203	5	55.6	8	2	Q70Y84_9LAMI	Q70y84 plectranthu
131	5	55.6	8	2	Q86BS9_STRPU	Q86BS9	strongyloce	204	5	55.6	8	2	Q7GEF5_NICBI	Q7gef5 nicotiana b
132	5	55.6	8	2	Q94623_MANSE	Q94623	manduca sex	205	5	55.6	8	2	Q7IMR4_BRANA	Q7imr4 brassica na
133	5	55.6	8	2	Q9N6M5_TOXGO	Q9N6M5	toxoplasma	206	5	55.6	8	2	Q7M1F1_RAPSA	Q7m1f1 raphanus sa
134	5	55.6	8	2	Q7GEM6_BRALA	Q7GEM6	brachioasto	207	5	55.6	8	2	Q7M267_GASCO	Q7m267 gastrocloni
135	5	55.6	8	2	Q8WGD7_9EUC	Q8WGD7	lomis hirta	208	5	55.6	8	2	Q7XB03_MAIZE	Q7xb03 zea mays (m
136	5	55.6	8	2	Q4YOC0_PLACH	Q4YOC0	plasmodium	209	5	55.6	8	2	Q8GTG5_LYCES	Q8gtg5 lycopersico
137	5	55.6	8	2	Q7M4A4_MERMC	Q7M4A4	mercenaria	210	5	55.6	8	2	Q8L802_MAIZE	Q8l802 zea mays (m
138	5	55.6	8	2	Q28866_MEGNO	Q28866	megaptera n	211	5	55.6	8	2	Q8MC14_9MYRT	Q8mc14 fuchsia hyb
139	5	55.6	8	2	Q5RC44_PONPY	Q5RC44	pongo pygma	212	5	55.6	8	2	Q8MC16_9MYRT	Q8mc16 ludwigia hy
140	5	55.6	8	2	Q5RLS9_PIG	Q5RLS9	sus scrofa	213	5	55.6	8	2	Q8MC18_QUIIN	Q8mc18 quibqualis
141	5	55.6	8	2	Q9BF92_TURTR	Q9BF92	tursiops tr	214	5	55.6	8	2	Q8MC20_9MYRT	Q8mc20 combretum w
142	5	55.6	8	2	Q9BF93_MEGNO	Q9BF93	megaptera n	215	5	55.6	8	2	Q8MC22_9MYRT	Q8mc22 sonneratia
143	5	55.6	8	2	Q9GMH3_LAGOB	Q9GMH3	lagenorhinc	216	5	55.6	8	2	Q8MC24_9MYRT	Q8mc24 sonneratia
144	5	55.6	8	2	Q70KG9_PIG	Q70KG9	sus scrofa	217	5	55.6	8	2	Q8MC26_9MYRT	Q8mc26 lagerstroem
145	5	55.6	8	2	Q4G3W0_MACMU	Q4G3W0	macaca mula	218	5	55.6	8	2	Q8MC28_9MYRT	Q8mc28 duabanga gr
146	5	55.6	8	2	Q37854_BPR17	Q37854	bacterioph	219	5	55.6	8	2	Q8MC30_LAGSP	Q8mc30 lagerstroem
147	5	55.6	8	2	Q5W4V4_9VIRU	Q5W4V4	bacterioph	220	5	55.6	8	2	Q8MC32_9MYRT	Q8mc32 rotala indi
148	5	55.6	8	2	Q5W4V6_9CAUD	Q5W4V6	bacterioph	221	5	55.6	8	2	Q8MC34_9MYRT	Q8mc34 heimia myrt
149	5	55.6	8	2	Q5W4V8_BPR51	Q5W4V8	bacterioph	222	5	55.6	8	2	Q8MC40_9MYRT	Q8mc40 nesaea lued
150	5	55.6	8	2	Q5W4W1_9CAUD	Q5W4W1	bacterioph	223	5	55.6	8	2	Q8MC42_9MYRT	Q8mc42 sonneratia
151	5	55.6	8	2	Q5W4W2_BPR32	Q5W4W2	bacterioph	224	5	55.6	8	2	Q8MC46_9MYRT	Q8mc46 peplis port
152	5	55.6	8	2	Q5W4W4_9CAUD	Q5W4W4	bacterioph	225	5	55.6	8	2	Q8MC48_9MYRT	Q8mc48 pemphis aci
153	5	55.6	8	2	Q5W4W6_9VIRU	Q5W4W6	bacterioph	226	5	55.6	8	2	Q8MC50_CUPLA	Q8mc50 cuphea lanc
154	5	55.6	8	2	Q5W4W8_9CAUD	Q5W4W8	bacterioph	227	5	55.6	8	2	Q8MC52_9MYRT	Q8mc52 woodfordia
155	5	55.6	8	2	Q5W4X0_9VIRU	Q5W4X0	bacterioph	228	5	55.6	8	2	Q8MES7_USRUD	Q8mes7 daphniphyll
156	5	55.6	8	2	Q5W4X2_9VIRU	Q5W4X2	bacterioph	229	5	55.6	8	2	Q8MES9_CERJA	Q8mes9 cercidiphyll
157	5	55.6	8	2	Q5W4X4_9VIRU	Q5W4X4	bacterioph	230	5	55.6	8	2	Q8MET1_USRUD	Q8met1 saxifraga s
158	5	55.6	8	2	Q5W4X6_BPL25	Q5W4X6	bacterioph	231	5	55.6	8	2	Q9GC24_9LILI	Q9gc24 nenga punil
159	5	55.6	8	2	Q5W4X8_BPL24	Q5W4X8	bacterioph	232	5	55.6	8	2	Q9GD00_9LILI	Q9gd00 masoala mad
160	5	55.6	8	2	Q5W4Y0_BPL21	Q5W4Y0	bacterioph	233	5	55.6	8	2	Q9GD47_9LILI	Q9gd47 hydriastele
161	5	55.6	8	2	Q5W4Y2_9CAUD	Q5W4Y2	bacterioph	234	5	55.6	8	2	Q56ZT0_ARATH	Q56zt0 arabidopsis
162	5	55.6	8	2	Q5W4Y4_9CAUD	Q5W4Y4	bacterioph	235	5	55.6	8	2	Q8LW32_9MYRT	Q8lw32 trapa maxim
163	5	55.6	8	2	Q5W4Y6_BPR70	Q5W4Y6	bacterioph	236	5	55.6	8	2	Q8LW33_PUNGR	Q8lw33 punica gran
164	5	55.6	8	2	Q6LDY8_BPF1	Q6LDY8	bacterioph	237	5	55.6	8	2	Q8LW34_LYRSA	Q8lw34 lythrum sal
165	5	55.6	8	2	Q8H9H3_9CAUD	Q8H9H3	bacterioph	238	5	55.6	8	2	Q8MC36_9MYRT	Q8mc36 lawsonia in
166	5	55.6	8	2	Q8H9H5_BPT6	Q8H9H5	bacterioph	239	5	55.6	8	2	Q8MC38_9MYRT	Q8mc38 ammannia ba
167	5	55.6	8	2	Q8H9H6_BPT2	Q8H9H6	bacterioph	240	5	55.6	8	2	Q8MC44_DECVE	Q8mc44 decodon ver
168	5	55.6	8	2	Q8H9T2_BPR27	Q8H9T2	bacterioph	241	5	55.6	8	2	Q4JLA6_BETVU	Q4jla6 beta vulgar
169	5	55.6	8	2	Q8H9T3_9CAUD	Q8H9T3	bacterioph	242	5	55.6	8	2	Q6ZZ01_SILCO	Q6zz01 silene conl
170	5	55.6	8	2	Q8H9T4_9CAUD	Q8H9T4	bacterioph	243	5	55.6	8	2	Q32560_ECOLI	Q32560 escherichia
171	5	55.6	8	2	Q8H9T6_BPR15	Q8H9T6	bacterioph	244	5	55.6	8	2	O52062_BACWE	O52062 bacillus me
172	5	55.6	8	2	Q8H9T8_9CAUD	Q8H9T8	bacterioph	245	5	55.6	8	2	P72279_RHOGO	P72279 rhodococcus
173	5	55.6	8	2	Q8H9T9_BPR10	Q8H9T9	bacterioph	246	5	55.6	8	2	Q45889_CLOBO	Q45889 clostridium
174	5	55.6	8	2	Q8H9J1_BPR06	Q8H9J1	bacterioph	247	5	55.6	8	2	Q51594_9ZZZZ	Q51594 plasamid col
175	5	55.6	8	2	Q8H9J3_BPR03	Q8H9J3	bacterioph	248	5	55.6	8	2	Q56140_STRTR	Q56140 streptococc
176	5	55.6	8	2	Q8H9J5_9CAUD	Q8H9J5	bacterioph	249	5	55.6	8	2	Q6DQ21_CLOBE	Q6dq21 clostridium
177	5	55.6	8	2	Q8H9J7_BPOX2	Q8H9J7	bacterioph	250	5	55.6	8	2	Q6LAA8_SHIFL	Q6laa8 shigella fl

251	5	55.6	8	2	Q6UNM3_VIBCH	Q6umM3	vibrio chol	324	5	55.6	8	2	Q94V88_9SAUR	Q94v88	varanus tri
252	5	55.6	8	2	Q71UF7_MORMO	Q71uf7	morganella	325	5	55.6	8	2	Q94V91_VARTI	Q94v91	varanus tim
253	5	55.6	8	2	Q71V47_PSEPU	Q71v47	pseudomonas	326	5	55.6	8	2	Q94VA0_9SAUR	Q94va0	varanus sem
254	5	55.6	8	2	Q799V9_BACSH	Q799v9	bacillus sp	327	5	55.6	8	2	Q94VA7_9SAUR	Q94va7	varanus sal
255	5	55.6	8	2	Q79CX6_MYXXA	Q79cx6	myxococcus	328	5	55.6	8	2	Q94VB2_9SAUR	Q94vb2	varanus sal
256	5	55.6	8	2	Q79AJ4_9GAMM	Q79aj4	acinetobact	329	5	55.6	8	2	Q94VB5_9SAUR	Q94vb5	varanus sal
257	5	55.6	8	2	Q7BW19_VIBCH	Q7bw19	vibrio chol	330	5	55.6	8	2	Q94VC1_VASRU	Q94vc1	varanus rud
258	5	55.6	8	2	Q7DKL7_STAAR	Q7dkl7	staphylococ	331	5	55.6	8	2	Q94VE4_VARML	Q94ve4	varanus mel
259	5	55.6	8	2	Q7MI94_ECOLI	Q7mi94	escherichia	332	5	55.6	8	2	Q94VF3_9SAUR	Q94vf3	varanus kei
260	5	55.6	8	2	Q8RJ10_STRCS	Q8rj10	streptomyce	333	5	55.6	8	2	Q94VF6_VARJO	Q94vf6	varanus job
261	5	55.6	8	2	Q934S4_THIPE	Q934s4	thiobacillu	334	5	55.6	8	2	Q94VF9_VARNB	Q94vf9	varanus ind
262	5	55.6	8	2	Q9R9C2_BORBU	Q9rc2	borrelia bu	335	5	55.6	8	2	Q94VJ4_VARNB	Q94vj4	varanus ben
263	5	55.6	8	2	Q9S6D5_ECOLI	Q9s6d5	escherichia	336	5	55.6	8	2	Q9P869_CHICK	Q9p869	gallus gall
264	5	55.6	8	2	Q9ZIE9_NEMPE	Q9zie9	neisseria m	337	5	55.6	8	2	Q9TD02_9SMEG	Q9td02	terranatos
265	5	55.6	8	2	Q092S8_SYNPM	Q092s8	synecococc	338	5	55.6	8	2	Q6VMC6_9PASS	Q6vmc6	serilophus
266	5	55.6	8	2	Q79AG6_ENTAG	Q79ag6	enterobacte	339	5	55.6	9	1	BRK1_RANNI	Q71k54	rana nigrom
267	5	55.6	8	2	Q62721_RAT	Q62721	rattus norv	340	5	55.6	9	1	BSP43_SERPL	Q8t2k9	serrattia pl
268	5	55.6	8	2	Q7M056_9MURI	Q7m056	mus sp. gen	341	5	55.6	9	1	CB22_SPIOL	Q9t2k9	spinacia ol
269	5	55.6	8	2	Q80WD6_MOUSE	Q80wd6	mus musculu	342	5	55.6	9	1	CB2B_SPIOL	Q9t2k9	spinacia ol
270	5	55.6	8	2	Q80XV8_9MURI	Q80xv8	rattus? sp.	343	5	55.6	9	1	CB2C_SPIOL	Q9t2k9	spinacia ol
271	5	55.6	8	2	Q8CJ03_MOUSE	Q8cj03	mus musculu	344	5	55.6	9	1	CONO_CONGE	P05486	conus geogr
272	5	55.6	8	2	Q8R4D8_MOUSE	Q8r4d8	mus musculu	345	5	55.6	9	1	CONO_CONST	P05487	conus stria
273	5	55.6	8	2	Q8R5M9_MOUSE	Q8r5m9	mus musculu	346	5	55.6	9	1	DNF1_LOCFI	P16339	locusta mig
274	5	55.6	8	2	Q99WH2_MOUSE	Q99wh2	mus musculu	347	5	55.6	9	1	FAR1_CALVO	P41856	calliphora
275	5	55.6	8	2	Q99P40_MOUSE	Q99p40	mus musculu	348	5	55.6	9	1	FAR2_CALVO	P41857	calliphora
276	5	55.6	8	2	Q9ERD2_MOUSE	Q9erd2	mus musculu	349	5	55.6	9	1	FAR2_PANRE	P41873	panagrellus
277	5	55.6	8	2	Q9ET16_MESAU	Q9et16	mesocricetu	350	5	55.6	9	1	FAR3_CALVO	P41858	calliphora
278	5	55.6	8	2	Q9ET17_MUSCR	Q9et17	mus catoli	351	5	55.6	9	1	FAR3_MACRS	P83276	macrobrachi
279	5	55.6	8	2	Q9ET18_MUSCR	Q9et18	mus spretus	352	5	55.6	9	1	FAR3_PENMO	P83318	penaeus mon
280	5	55.6	8	2	Q9JLD7_MESAU	Q9jld7	mesocricetu	353	5	55.6	9	1	FAR4_CALVO	P83318	penaeus mon
281	5	55.6	8	2	Q9QVJ8_9MURI	Q9qvj8	mus sp. mep	354	5	55.6	9	1	FAR4_PENMO	P83319	penaeus mon
282	5	55.6	8	2	Q78ED1_RAT	Q78ed1	rattus norv	355	5	55.6	9	1	FAR5_ASCSU	P43170	ascaris suu
283	5	55.6	8	2	Q80WD5_MUSJP	Q80wd5	mus spretus	356	5	55.6	9	1	FAR5_CALVO	P41860	calliphora
284	5	55.6	8	2	Q89965_POVJC	Q89965	polymaviru	357	5	55.6	9	1	FAR5_PANRE	P82661	panagrellus
285	5	55.6	8	2	Q6PUD5_SV40	Q6pud5	simian viru	358	5	55.6	9	1	FAR5_PENMO	P83320	penaeus mon
286	5	55.6	8	2	Q6PUD7_SV40	Q6pud7	simian viru	359	5	55.6	9	1	FAR6_CALVO	P41861	calliphora
287	5	55.6	8	2	Q6PUD9_SV40	Q6pud9	simian viru	360	5	55.6	9	1	FAR6_MACRS	P83279	macrobrachi
288	5	55.6	8	2	Q6PUE1_SV40	Q6pue1	simian viru	361	5	55.6	9	1	FAR7_CALVO	P41862	calliphora
289	5	55.6	8	2	Q6QEX9_SV40	Q6qex9	simian viru	362	5	55.6	9	1	FAR8_MACRS	P83281	macrobrachi
290	5	55.6	8	2	Q7T863_9VIRU	Q7t863	largemouth	363	5	55.6	9	1	FAR9_ASCSU	P43172	ascaris suu
291	5	55.6	8	2	Q80H91_9PARA	Q80h91	newcastle d	364	5	55.6	9	1	FARA_CALVO	P41865	calliphora
292	5	55.6	8	2	Q83977_9INFA	Q83977	influenza a	365	5	55.6	9	1	FARD_CALVO	P41868	calliphora
293	5	55.6	8	2	Q84156_9POXV	Q84156	orf virus	366	5	55.6	9	1	FARD_CALSI	P38495	calinectes
294	5	55.6	8	2	Q84271_HPV19	Q84271	human papil	367	5	55.6	9	1	FIBB_ERYPA	P19346	erythrocebu
295	5	55.6	8	2	Q84273_HPV25	Q84273	human papil	368	5	55.6	9	1	FIBB_MACFU	P19345	macaca fusc
296	5	55.6	8	2	Q9DSN0_9VIRU	Q9dsn0	beet soil-b	369	5	55.6	9	1	FIBB_PAPAN	P19344	papio anubi
297	5	55.6	8	2	Q9DSN1_9VIRU	Q9dsn1	beet soil-b	370	5	55.6	9	1	FIBB_PAPHA	P19343	papio hamad
298	5	55.6	8	2	Q9DSN2_9VIRU	Q9dsn2	beet soil-b	371	5	55.6	9	1	FIBB_THERG	P19342	theropithec
299	5	55.6	8	2	Q9DSN3_9VIRU	Q9dsn3	beet soil-b	372	5	55.6	9	1	FMRP1_SARBU	P83350	sarcophaga
300	5	55.6	8	2	Q9DSN4_9VIRU	Q9dsn4	beet soil-b	373	5	55.6	9	1	HUTU_KLEAE	P12381	klebsiella
301	5	55.6	8	2	Q9DSN5_9VIRU	Q9dsn5	beet soil-b	374	5	55.6	9	1	IPYR_RHOVI	P82992	rhodopseudo
302	5	55.6	8	2	Q9DSN6_9VIRU	Q9dsn6	beet soil-b	375	5	55.6	9	1	KNL3_BOMVA	P83058	bombina var
303	5	55.6	8	2	Q9E8P7_9VIRU	Q9ep7	beet soil-b	376	5	55.6	9	1	KNL3_CYPDO	P83659	cophononyx
304	5	55.6	8	2	Q9E8P8_9VIRU	Q9ep8	beet soil-b	377	5	55.6	9	1	LMT3_LOCFI	P41889	locusta mig
305	5	55.6	8	2	Q9E8P9_9VIRU	Q9ep9	beet soil-b	378	5	55.6	9	1	NEUU_CAVPO	P34966	cavia porce
306	5	55.6	8	2	Q9E8Q0_9VIRU	Q9eq0	beet soil-b	379	5	55.6	9	1	NSKI_SARBU	P41492	sarcophaga
307	5	55.6	8	2	Q9E8Q1_9VIRU	Q9eq1	beet soil-b	380	5	55.6	9	1	OXYT_CYPCA	P69128	cypripus ca
308	5	55.6	8	2	Q9E8Q2_9VIRU	Q9eq2	beet soil-b	381	5	55.6	9	1	OXYT_BISFO	P42998	eisenia foe
309	5	55.6	8	2	Q9E8Q3_9VIRU	Q9eq3	beet soil-b	382	5	55.6	9	1	OXYT_OCTVU	P80027	octopus vul
310	5	55.6	8	2	Q9E8Q4_9VIRU	Q9eq4	beet soil-b	383	5	55.6	9	1	OXYT_PETMA	P69129	petromyzon
311	5	55.6	8	2	Q9E8Q5_9VIRU	Q9eq5	beet soil-b	384	5	55.6	9	1	PGLR_DIAAB	P81179	diaprepes a
312	5	55.6	8	2	Q9E8Q6_9VIRU	Q9eq6	beet soil-b	385	5	55.6	9	1	PPH1_LYCES	P83380	lycopersico
313	5	55.6	8	2	Q9E8Q7_9VIRU	Q9eq7	beet soil-b	386	5	55.6	9	1	PPK1_PERAM	P82691	periplaneta
314	5	55.6	8	2	Q9J205_9HEPC	Q9j205	hepatitis c	387	5	55.6	9	1	PKV2_MUSDO	P84355	musca domes
315	5	55.6	8	2	Q56T42_9GEMI	Q56t42	okra yellow	388	5	55.6	9	1	PKV2_SARBU	P84353	sarcophaga
316	5	55.6	8	2	P79940_XENLA	P79940	xenopus lae	389	5	55.6	9	1	PFY2_PENMO	P84006	penaeus mon
317	5	55.6	8	2	Q5YDB9_9PERC	Q5ydb9	xiphister m	390	5	55.6	9	1	PFY3_PENMO	P84007	penaeus mon
318	5	55.6	8	2	Q5YDW3_9PERC	Q5ydw3	xiphister m	391	5	55.6	9	1	PFY4_PENMO	P84008	penaeus mon
319	5	55.6	8	2	Q64IX4_FUNHE	Q64ix4	fundulus he	392	5	55.6	9	1	PFY LOLVU	P42004	loligo vulg
320	5	55.6	8	2	Q715L5_VARDU	Q715l5	varanus dum	393	5	55.6	9	1	RS10_SERMA	Q68936	serrattia ma
321	5	55.6	8	2	Q71Z46_CTEID	Q71z46	ctenopharyn	394	5	55.6	9	1	TKC1_CALVO	P41517	calliphora
322	5	55.6	8	2	Q8JFN8_CHICK	Q8jfn8	gallus gall	395	5	55.6	9	1	TKU1_LOCFI	P16223	locusta mig
323	5	55.6	8	2	Q94V82_9SAUR	Q94v82	varanus yuw	396	5	55.6	9	1	TRP4_LEUMA	P81736	leucophaea

397	5	55.6	9	1	ULAK_MOUSE	P99031_mus musculus	470	5	55.6	9	2	Q7JIT1_LAGAC	Q7jit1_lagenorhync
398	5	55.6	9	1	UN19_CLOPA	P81355_clostridium	471	5	55.6	9	2	Q7M375_BOVIN	Q7m375_bos taurus
399	5	55.6	9	1	XYLA_STRS8	P19149_streptomyces	472	5	55.6	9	2	Q8MJN1_CEBPY	Q8mjn1_cebuela py
400	5	55.6	9	1	YBFR_AZOV1	P25825_azotobacter	473	5	55.6	9	2	Q8MJN2_CALJA	Q8mjn2_calithrix
401	5	55.6	9	2	Q50832_METVO	P25832_methanococc	474	5	55.6	9	2	Q8MJN3_CALGO	Q8mjn3_calimico g
402	5	55.6	9	2	Q7R899_NEUCR	Q7ty89_neurospora	475	5	55.6	9	2	Q8MJN4_LEORO	Q8mjn4_leontopithe
403	5	55.6	9	2	Q7S182_NEUCR	Q7s182_neurospora	476	5	55.6	9	2	Q8MJN5_SAGFU	Q8mjn5_saguinus fu
404	5	55.6	9	2	Q7SCD2_NEUCR	Q7sacd2_neurospora	477	5	55.6	9	2	Q8MJN6_AOTAZ	Q8mjn6_aotus azara
405	5	55.6	9	2	Q7UR18_9ASCO	Q9ur18_sclerotium	478	5	55.6	9	2	Q8MJN7_SAISC	Q8mjn7_saimiri sci
406	5	55.6	9	2	Q523Q3_MAGGR	Q523q3_magnaporthe	479	5	55.6	9	2	Q8MJN8_CEBAP	Q8mjn8_cebus apell
407	5	55.6	9	2	Q14277_HUMAN	Q14277_homo sapien	480	5	55.6	9	2	Q8MJN9_ATEFS	Q8mjn9_ateles fusc
408	5	55.6	9	2	Q15892_HUMAN	Q15892_homo sapien	481	5	55.6	9	2	Q9GJVI_LAGAC	Q9gji1_lagenorhync
409	5	55.6	9	2	Q16220_HUMAN	Q16220_homo sapien	482	5	55.6	9	2	Q9GJVI_LAGOL	Q9gji2_lagenorhync
410	5	55.6	9	2	Q16276_HUMAN	Q16276_homo sapien	483	5	55.6	9	2	Q9GJVI_LAGOB	Q9gji3_lagenorhync
411	5	55.6	9	2	Q16386_HUMAN	Q16386_homo sapien	484	5	55.6	9	2	Q9TRW2_RABIT	Q9trw2_oryctolagus
412	5	55.6	9	2	Q67AQ6_HUMAN	Q67aq6_homo sapien	485	5	55.6	9	2	Q9TIT7_BOVIN	Q9tit7_bos taurus
413	5	55.6	9	2	Q67AQ7_HUMAN	Q67aq7_homo sapien	486	5	55.6	9	2	Q9TUY0_MONDO	Q9tuy0_monodelphis
414	5	55.6	9	2	Q67AQ8_HUMAN	Q67aq8_homo sapien	487	5	55.6	9	2	Q9XT05_MACRG	Q9xt05_macropus ru
415	5	55.6	9	2	Q67AR0_HUMAN	Q67ar0_homo sapien	488	5	55.6	9	2	Q6JDL5_CANFA	Q6jdl5_canis fami1
416	5	55.6	9	2	Q67AR1_HUMAN	Q67ar1_homo sapien	489	5	55.6	9	2	Q4TZV5_PAPHA	Q4tzv5_papio hamad
417	5	55.6	9	2	Q67AR4_HUMAN	Q67ar4_homo sapien	490	5	55.6	9	2	Q38366_BPHX	Q38366_bacterioph
418	5	55.6	9	2	Q67AR5_HUMAN	Q67ar5_homo sapien	491	5	55.6	9	2	Q42452_WHEAT	Q42452_criticum ae
419	5	55.6	9	2	Q67AR6_HUMAN	Q67ar6_homo sapien	492	5	55.6	9	2	Q5D4X3_9MYRT	Q5d4x3_lourtellia r
420	5	55.6	9	2	Q67AR7_HUMAN	Q67ar7_homo sapien	493	5	55.6	9	2	Q5D4Y6_9MYRT	Q5d4y6_lythrum hys
421	5	55.6	9	2	Q67AS0_HUMAN	Q67as0_homo sapien	494	5	55.6	9	2	Q5D4Y6_9MYRT	Q5d4y6_capuronia m
422	5	55.6	9	2	Q67AS3_HUMAN	Q67as3_homo sapien	495	5	55.6	9	2	Q5IA44_9ARAE	Q5ia44_philodendro
423	5	55.6	9	2	Q67AT1_HUMAN	Q67at1_homo sapien	496	5	55.6	9	2	Q5IA45_MAHBE	Q5ia45_mahonia bea
424	5	55.6	9	2	Q67AT2_HUMAN	Q67at2_homo sapien	497	5	55.6	9	2	Q5VB50_9FILI	Q5vb50_boibitis au
425	5	55.6	9	2	Q67AT6_HUMAN	Q67at6_homo sapien	498	5	55.6	9	2	Q5VB51_9FILI	Q5vb51_elaphogloss
426	5	55.6	9	2	Q6KER0_HUMAN	Q6ker0_homo sapien	499	5	55.6	9	2	Q5VB52_9FILI	Q5vb52_elaphogloss
427	5	55.6	9	2	Q6LCV2_HUMAN	Q6lcv2_homo sapien	500	5	55.6	9	2	Q5VB53_9FILI	Q5vb53_elaphogloss
428	5	55.6	9	2	Q6LEH2_HUMAN	Q6leh2_homo sapien	501	5	55.6	9	2	Q5VB54_9FILI	Q5vb54_elaphogloss
429	5	55.6	9	2	Q71KUP3_HUMAN	Q71kup3_homo sapien	502	5	55.6	9	2	Q5VB55_9FILI	Q5vb55_elaphogloss
430	5	55.6	9	2	Q7LYR3_HUMAN	Q7lyr3_homo sapien	503	5	55.6	9	2	Q5VB56_9FILI	Q5vb56_elaphogloss
431	5	55.6	9	2	Q7KYP6_HUMAN	Q7kyp6_homo sapien	504	5	55.6	9	2	Q5VB57_9FILI	Q5vb57_elaphogloss
432	5	55.6	9	2	Q7M4S2_HUMAN	Q7m4s2_homo sapien	505	5	55.6	9	2	Q5VB58_9FILI	Q5vb58_elaphogloss
433	5	55.6	9	2	Q7Z4P0_HUMAN	Q7z4p0_homo sapien	506	5	55.6	9	2	Q5VB59_9FILI	Q5vb59_elaphogloss
434	5	55.6	9	2	Q9BYF9_HUMAN	Q9byf9_homo sapien	507	5	55.6	9	2	Q5VB60_9FILI	Q5vb60_elaphogloss
435	5	55.6	9	2	Q9H4M8_HUMAN	Q9h4m8_homo sapien	508	5	55.6	9	2	Q5VB61_9FILI	Q5vb61_elaphogloss
436	5	55.6	9	2	Q9UC36_HUMAN	Q9uc36_homo sapien	509	5	55.6	9	2	Q5VB62_9FILI	Q5vb62_elaphogloss
437	5	55.6	9	2	Q9UE26_HUMAN	Q9ue26_homo sapien	510	5	55.6	9	2	Q5VB63_9FILI	Q5vb63_elaphogloss
438	5	55.6	9	2	Q9UKJ6_HUMAN	Q9ukj6_homo sapien	511	5	55.6	9	2	Q5VB64_9FILI	Q5vb64_elaphogloss
439	5	55.6	9	2	Q9UNAO_HUMAN	Q9unao_homo sapien	512	5	55.6	9	2	Q5VB65_9FILI	Q5vb65_elaphogloss
440	5	55.6	9	2	Q9UQA3_HUMAN	Q9uqa3_homo sapien	513	5	55.6	9	2	Q5VB66_9FILI	Q5vb66_elaphogloss
441	5	55.6	9	2	Q14715_HUMAN	Q14715_homo sapien	514	5	55.6	9	2	Q5VB67_9FILI	Q5vb67_elaphogloss
442	5	55.6	9	2	Q15999_HUMAN	Q15999_homo sapien	515	5	55.6	9	2	Q5VB68_9FILI	Q5vb68_elaphogloss
443	5	55.6	9	2	Q27396_BABBO	Q27396_babesia bov	516	5	55.6	9	2	Q5VB69_9FILI	Q5vb69_elaphogloss
444	5	55.6	9	2	O5C1F8_SCHJA	O5cif8_schistosoma	517	5	55.6	9	2	Q5VB70_9FILI	Q5vb70_elaphogloss
445	5	55.6	9	2	Q7JNB6_DROVI	Q7jnb6_drosophila	518	5	55.6	9	2	Q5VB71_9FILI	Q5vb71_elaphogloss
446	5	55.6	9	2	Q7M3L3_PENVA	Q7m3l3_penaeus van	519	5	55.6	9	2	Q5VB72_9FILI	Q5vb72_elaphogloss
447	5	55.6	9	2	Q7M3N6_GRYBI	Q7m3n6_gryllus bim	520	5	55.6	9	2	Q5VB73_9FILI	Q5vb73_elaphogloss
448	5	55.6	9	2	Q7M3N7_GRYBI	Q7m3n7_gryllus bim	521	5	55.6	9	2	Q5VB74_9FILI	Q5vb74_elaphogloss
449	5	55.6	9	2	Q7M3N8_GRYBI	Q7m3n8_gryllus bim	522	5	55.6	9	2	Q5VB75_9FILI	Q5vb75_elaphogloss
450	5	55.6	9	2	Q9TDW6_LEPDE	Q9tdw6_leptinotars	523	5	55.6	9	2	Q5VB76_9FILI	Q5vb76_elaphogloss
451	5	55.6	9	2	Q9TWV0_ATEL	Q9twv0_anthopleura	524	5	55.6	9	2	Q5VB77_9FILI	Q5vb77_elaphogloss
452	5	55.6	9	2	Q9TWX7_MANSE	Q9twx7_manduca sex	525	5	55.6	9	2	Q5VB78_9FILI	Q5vb78_elaphogloss
453	5	55.6	9	2	P84502_9ANNE	P84502_annelida. 1	526	5	55.6	9	2	Q5VB79_9FILI	Q5vb79_elaphogloss
454	5	55.6	9	2	Q7RE58_PLAYO	Q7res8_plasmodium	527	5	55.6	9	2	Q5VB80_9FILI	Q5vb80_elaphogloss
455	5	55.6	9	2	Q7RGW1_PLAYO	Q7rgw1_plasmodium	528	5	55.6	9	2	Q5VB81_9FILI	Q5vb81_elaphogloss
456	5	55.6	9	2	Q8WGE6_PROCL	Q8wge6_procambarus	529	5	55.6	9	2	Q5VB82_9FILI	Q5vb82_elaphogloss
457	5	55.6	9	2	Q4X912_PLACH	Q4x912_plasmodium	530	5	55.6	9	2	Q5VB83_9FILI	Q5vb83_elaphogloss
458	5	55.6	9	2	Q4Y413_PLACH	Q4y413_plasmodium	531	5	55.6	9	2	Q5VB84_9FILI	Q5vb84_elaphogloss
459	5	55.6	9	2	Q4Y954_PLACH	Q4y954_plasmodium	532	5	55.6	9	2	Q5VB85_9FILI	Q5vb85_elaphogloss
460	5	55.6	9	2	Q4Y9N3_PLABE	Q4y9n3_plasmodium	533	5	55.6	9	2	Q5VB86_9FILI	Q5vb86_elaphogloss
461	5	55.6	9	2	Q6LDS7_RABIT	Q6lds7_oryctolagus	534	5	55.6	9	2	Q5VB87_9FILI	Q5vb87_elaphogloss
462	5	55.6	9	2	Q7JIS3_LAGOB	Q7jis3_lagenorhync	535	5	55.6	9	2	Q5VB88_9FILI	Q5vb88_elaphogloss
463	5	55.6	9	2	Q7JIS4_LAGOL	Q7jis4_lagenorhync	536	5	55.6	9	2	Q5VB89_9FILI	Q5vb89_elaphogloss
464	5	55.6	9	2	Q7JIS5_LAGOL	Q7jis5_lagenorhync	537	5	55.6	9	2	Q5VB90_9FILI	Q5vb90_elaphogloss
465	5	55.6	9	2	Q7JIS6_LAGOL	Q7jis6_lagenorhync	538	5	55.6	9	2	Q5VB91_9FILI	Q5vb91_elaphogloss
466	5	55.6	9	2	Q7JIS7_LAGOL	Q7jis7_lagenorhync	539	5	55.6	9	2	Q5VB92_9FILI	Q5vb92_elaphogloss
467	5	55.6	9	2	Q7JIS8_LAGOL	Q7jis8_lagenorhync	540	5	55.6	9	2	Q5VB93_9FILI	Q5vb93_elaphogloss
468	5	55.6	9	2	Q7JIS9_LAGAC	Q7jis9_lagenorhync	541	5	55.6	9	2	Q5VB94_9FILI	Q5vb94_elaphogloss
469	5	55.6	9	2	Q7JIT0_LAGAC	Q7jit0_lagenorhync	542	5	55.6	9	2	Q5VB95_9FILI	Q5vb95_elaphogloss



543	5	55.6	9	2	Q5VB96_9FILI	Q5vb96 elaphogloss	616	5	55.6	9	2	Q5K2V7_NODSP	Q5k2v7 nodularia s
544	5	55.6	9	2	Q5VB97_9FILI	Q5vb97 elaphogloss	617	5	55.6	9	2	P82568_STRPY	P82568 streptococc
545	5	55.6	9	2	Q5VB98_9FILI	Q5vb98 elaphogloss	618	5	55.6	9	2	P83539_LACSN	P83539 lactobacilli
546	5	55.6	9	2	Q5VB99_9FILI	Q5vb99 elaphogloss	619	5	55.6	9	2	Q6LDL7_STAAU	Q6ldl7 streptomyce
547	5	55.6	9	2	Q5VBA0_9FILI	Q5vba0 elaphogloss	620	5	55.6	9	2	Q6VCX0_9ACTO	Q6vcx0 vibrio fisc
548	5	55.6	9	2	Q5VBA1_9FILI	Q5vba1 elaphogloss	621	5	55.6	9	2	Q6VFO2_VIBFI	Q6vfo2 vibrio fisc
549	5	55.6	9	2	Q5VBA2_9FILI	Q5vba2 elaphogloss	622	5	55.6	9	2	Q711U9_9RHIZ	Q711u9 azorhizobiu
550	5	55.6	9	2	Q5VBA3_9FILI	Q5vba3 elaphogloss	623	5	55.6	9	2	Q711V3_AZOCA	Q711v3 azorhizobiu
551	5	55.6	9	2	Q5VBA4_9FILI	Q5vba4 elaphogloss	624	5	55.6	9	2	Q712B7_SYNTE	Q712b7 sinorhizobi
552	5	55.6	9	2	Q5VBA5_9FILI	Q5vba5 elaphogloss	625	5	55.6	9	2	Q71UF6_YERIN	Q71uf6 yerinia en
553	5	55.6	9	2	Q5VBA6_9FILI	Q5vba6 elaphogloss	626	5	55.6	9	2	Q798K5_STRLI	Q798k5 streptomyce
554	5	55.6	9	2	Q5VBA7_9FILI	Q5vba7 elaphogloss	627	5	55.6	9	2	Q7MOL7_9STRE	Q7mol7 streptococc
555	5	55.6	9	2	Q5VBA8_9FILI	Q5vba8 elaphogloss	628	5	55.6	9	2	Q7M151_9BACT	Q7m151 unidentified
556	5	55.6	9	2	Q5VBA9_9FILI	Q5vba9 elaphogloss	629	5	55.6	9	2	Q8GI26_LACDL	Q8gi26 lactobacilli
557	5	55.6	9	2	Q5VBB0_9FILI	Q5vbb0 elaphogloss	630	5	55.6	9	2	Q8RKU3_BORBU	Q8rku3 borrelia bu
558	5	55.6	9	2	Q5VBB1_9FILI	Q5vbb1 elaphogloss	631	5	55.6	9	2	Q93E20_STRAG	Q93e20 streptococc
559	5	55.6	9	2	Q5VBB2_9FILI	Q5vbb2 elaphogloss	632	5	55.6	9	2	Q9JN16_STRPY	Q9jni6 streptococc
560	5	55.6	9	2	Q5VBB3_9FILI	Q5vbb3 elaphogloss	633	5	55.6	9	2	Q9R7H9_HABIN	Q9r7h9 haemophilus
561	5	55.6	9	2	Q5VBB4_9FILI	Q5vbb4 elaphogloss	634	5	55.6	9	2	Q51765_PSEFL	Q51765 pseudomonas
562	5	55.6	9	2	Q5VBB5_9FILI	Q5vbb5 elaphogloss	635	5	55.6	9	2	Q08979_MOUSE	Q08979 mus musculus
563	5	55.6	9	2	Q5VBB6_9FILI	Q5vbb6 elaphogloss	636	5	55.6	9	2	Q6LQ11_RAT	Q6lq11 rattus norv
564	5	55.6	9	2	Q5VBB7_9FILI	Q5vbb7 elaphogloss	637	5	55.6	9	2	Q6YF34_RAT	Q6yfg34 rattus norv
565	5	55.6	9	2	Q5VBB8_9FILI	Q5vbb8 elaphogloss	638	5	55.6	9	2	Q78E72_RAT	Q78e72 rattus norv
566	5	55.6	9	2	Q5VBB9_9FILI	Q5vbb9 elaphogloss	639	5	55.6	9	2	Q7M078_RAT	Q7m078 rattus norv
567	5	55.6	9	2	Q5VBC0_9FILI	Q5vbc0 elaphogloss	640	5	55.6	9	2	Q8R514_RAT	Q8r514 rattus norv
568	5	55.6	9	2	Q5VBC1_9FILI	Q5vbc1 elaphogloss	641	5	55.6	9	2	Q92012_MOUSE	Q92012 mus musculus
569	5	55.6	9	2	Q5VBC2_9FILI	Q5vbc2 elaphogloss	642	5	55.6	9	2	Q9QVH9_9MURI	Q9qvh9 mus sp. sup
570	5	55.6	9	2	Q5VBC3_9FILI	Q5vbc3 elaphogloss	643	5	55.6	9	2	Q9QZA7_MOUSE	Q9qza7 mus musculus
571	5	55.6	9	2	Q5VBC4_9FILI	Q5vbc4 elaphogloss	644	5	55.6	9	2	Q9QWG2_MOUSE	Q9qwg2 mus musculus
572	5	55.6	9	2	Q5VBC5_9FILI	Q5vbc5 elaphogloss	645	5	55.6	9	2	Q88889_MOUSE	Q88889 mus musculus
573	5	55.6	9	2	Q5VBC6_9FILI	Q5vbc6 elaphogloss	646	5	55.6	9	2	Q61723_MOUSE	Q61723 mus musculus
574	5	55.6	9	2	Q5VBC7_9FILI	Q5vbc7 elaphogloss	647	5	55.6	9	2	Q62530_MUSSP	Q62530 mus spretus
575	5	55.6	9	2	Q5VBC8_9FILI	Q5vbc8 elaphogloss	648	5	55.6	9	2	Q4W8Q9_MOUSE	Q4w8q9 mus musculus
576	5	55.6	9	2	Q5VBC9_9FILI	Q5vbc9 elaphogloss	649	5	55.6	9	2	Q71066_9PARA	Q71066 canine dist
577	5	55.6	9	2	Q5Y9F5_9BRYO	Q5y9f5 calliegon	650	5	55.6	9	2	Q71067_9PARA	Q71067 canine dist
578	5	55.6	9	2	Q67218_PENAM	Q67218 pennisetum	651	5	55.6	9	2	Q71068_9PARA	Q71068 canine dist
579	5	55.6	9	2	Q6ALH7_HORVD	Q6alh7 hordeum vul	652	5	55.6	9	2	P90359_9POTY	P90359 barley mild
580	5	55.6	9	2	Q6EUV8_GERHY	Q6euv8 gerbera hyb	653	5	55.6	9	2	Q66545_9GAMA	Q66545 human herpe
581	5	55.6	9	2	Q6EX64_9LAMI	Q6ex64 hyptis flor	654	5	55.6	9	2	Q67605_SLCV	Q67605 squash leaf
582	5	55.6	9	2	Q6RVW6_CAPAN	Q6rvw6 capsicum an	655	5	55.6	9	2	Q67606_SLCV	Q67606 squash leaf
583	5	55.6	9	2	Q6VR25_9THECC	Q6vr25 theobroma c	656	5	55.6	9	2	Q69349_HLV2	Q69349 human herpe
584	5	55.6	9	2	Q6XBN2_9BRYO	Q6xbn2 vittia pach	657	5	55.6	9	2	Q69473_HVNI	Q69473 human herpe
585	5	55.6	9	2	Q6ZZ00_9CONTO	Q6zz00 silene rotu	658	5	55.6	9	2	Q6TIE2_CVHSA	Q6tie2 sars corona
586	5	55.6	9	2	Q70Y63_9CONTO	Q70y63 congea tome	659	5	55.6	9	2	Q82622_9COCO	Q82622 avian infec
587	5	55.6	9	2	Q70Y76_9SOLSC	Q70y76 solenotemo	660	5	55.6	9	2	Q83622_9FLAV	Q83622 murray vall
588	5	55.6	9	2	Q70Y83_9LAMI	Q70y83 electranthu	661	5	55.6	9	2	Q84333_9VAV	Q84333 simian viru
589	5	55.6	9	2	Q7EXPD_9HORVD	Q7expd hordeum vul	662	5	55.6	9	2	Q88953_9POXV	Q88953 vaccinia vi
590	5	55.6	9	2	Q7X8P7_MAYZE	Q7x8p7 zea mays (m	663	5	55.6	9	2	Q8QVD3_9MONO	Q8qvd3 ovine respi
591	5	55.6	9	2	Q85G96_9BRYO	Q85g96 pyrrhobryum	664	5	55.6	9	2	Q9ELU7_HPBVO	Q9elu7 hepatitis b
592	5	55.6	9	2	Q85V64_EUCGR	Q85v64 eucalyptus	665	5	55.6	9	2	Q9IBM8_9POLY	Q9ibm8 simian viru
593	5	55.6	9	2	Q8MBF4_IPOQU	Q8mbf4 ipomoea qua	666	5	55.6	9	2	Q9PYK1_9POLY	Q9pyk1 simian viru
594	5	55.6	9	2	Q8MEM3_9ROSI	Q8mem3 howittia tr	667	5	55.6	9	2	P84497_TRASC	P84497 trachemys s
595	5	55.6	9	2	Q94OK4_9BATH	Q94ok4 arabidopsis	668	5	55.6	9	2	Q5YDVL_9PRRC	Q5ydv1 xiphister m
596	5	55.6	9	2	Q95GN1_PELHO	Q95gn1 pelargonium	669	5	55.6	9	2	Q64IX3_FUNHE	Q64ix3 fundulus he
597	5	55.6	9	2	Q9AXH8_MESCR	Q9axh8 mesembryant	670	5	55.6	9	2	Q673X5_9CORV	Q673x5 platysteira
598	5	55.6	9	2	Q9FEC0_HORVU	Q9fec0 hordeum vul	671	5	55.6	9	2	Q673Y1_9CORV	Q673y1 lanoturdus
599	5	55.6	9	2	Q9GCV6_9LILI	Q9gcv6 sclerosperm	672	5	55.6	9	2	Q673Y6_9CORV	Q673y6 platysteira
600	5	55.6	9	2	Q9GDL2_9LILI	Q9gdl2 linospadix	673	5	55.6	9	2	Q673Z9_9CORV	Q673z9 batis poens
601	5	55.6	9	2	Q9S8J8_ORYSA	Q9s8j8 oryza fativ	674	5	55.6	9	2	Q691D6_ANOSA	Q691d6 anolis sagr
602	5	55.6	9	2	Q9TKD9_9MYRT	Q9tkd9 pericalymma	675	5	55.6	9	2	Q6HA69_TRITG	Q6ha69 trimeresuru
603	5	55.6	9	2	Q9TKF2_9MYRT	Q9tkf2 asteromyritu	676	5	55.6	9	2	Q6HA76_9SAUR	Q6ha76 trimeresuru
604	5	55.6	9	2	Q9TKG1_9MYRT	Q9tkg1 calothammus	677	5	55.6	9	2	Q71DX2_9SAUR	Q71dx2 urostrophus
605	5	55.6	9	2	Q9TLM7_9FLOR	Q9tlm7 laurentia v	678	5	55.6	9	2	Q75QC9_GALLIA	Q75qc9 gallus lafa
606	5	55.6	9	2	Q5SBX0_9POAL	Q5sbx0 restio insi	679	5	55.6	9	2	Q71Z50_CHICK	Q71z50 gallus gall
607	5	55.6	9	2	Q4QWV3_9MARC	Q4qwv3 pallavicini	680	5	55.6	9	2	Q71Z66_MEUGA	Q71z66 meleagris g
608	5	55.6	9	2	Q4QWV9_9MARC	Q4qwv9 jensenia sp	681	5	55.6	9	2	Q71ZJ8_RANTE	Q71zj8 rana tempor
609	5	55.6	9	2	Q4U0F2_MALDO	Q4u0f2 malus domes	682	5	55.6	9	2	Q8AYL5_CAFARU	Q8ayl5 carassius a
610	5	55.6	9	2	Q3U0790_ERWAM	Q3u0790 erwinia amy	683	5	55.6	9	2	Q8SHF0_CHANA	Q8shf0 chamaeleo n
611	5	55.6	9	2	P72345_PSEEX	P72345 pseudomonas	684	5	55.6	9	2	Q92009_CHICK	Q92009 gallus gall
612	5	55.6	9	2	Q45852_CLOSTR	Q45852 clostridium	685	5	55.6	9	2	Q94VC6_9SAUR	Q94vc6 varanus nil
613	5	55.6	9	2	Q47410_ECOLI	Q47410 escherichia	686	5	55.6	9	2	Q94VD8_VARNI	Q94vd8 varanus nil
614	5	55.6	9	2	Q53914_STREPTO	Q53914 streptomyce	687	5	55.6	9	2	Q94VE1_VARME	Q94ve1 varanus mer
615	5	55.6	9	2	Q5K2V4_9NOST	Q5k2v4 nodularia h	688	5	55.6	9	2	Q94VG2_VARIN	Q94vg2 varanus ind

689	5	55.6	9	2	Q94VH4_9SAUR	Q94VH4	varanus gla	762	5	55.6	10	1	UPA9_HUMAN	P30095	homo sapien
690	5	55.6	9	2	Q94V10_VARGI	Q94V10	varanus gig	763	5	55.6	10	2	Q7M530_PYRFU	Q7M530	pyrococcus
691	5	55.6	9	2	Q94V18_VARER	Q94V18	varanus ere	764	5	55.6	10	2	Q7M4X7_FUSSP	Q7M4X7	fusarium sp
692	5	55.6	9	2	Q94VJ1_VARDOR	Q94VJ1	varanus dor	765	5	55.6	10	2	Q7RYS0_NEUCR	Q7RYS0	neurospora
693	5	55.6	9	2	Q9TAI4_CHICK	Q9TAI4	gallus gall	766	5	55.6	10	2	Q7RZZ2_NEUCR	Q7RZZ2	neurospora
694	5	55.6	9	2	Q9PRJ4_LEPOS	Q9PRJ4	lepisosteus	767	5	55.6	10	2	Q7S184_NEUCR	Q7S184	neurospora
695	5	55.6	9	2	Q9T688_GEGCE	Q9T688	gecko gecko	768	5	55.6	10	2	Q7S377_NEUCR	Q7S377	neurospora
696	5	55.6	9	2	Q5UGN1_CHICK	Q5UGN1	gallus gall	769	5	55.6	10	2	Q7S5J5_NEUCR	Q7S5J5	neurospora
697	5	55.6	9	2	Q4PU39_9CICH	Q4PU39	lepidolamp	770	5	55.6	10	2	Q7SA62_NEUCR	Q7SA62	neurospora
698	5	55.6	9	2	Q7LZ17_9NEOB	Q7LZ17	heleophryne	771	5	55.6	10	2	Q9C1K4_GLOMO	Q9C1K4	glomus moss
699	5	55.6	9	2	Q12096_CAEV	Q12096	caprine art	772	5	55.6	10	2	Q1S342_HUMAN	Q1S342	homo sapien
700	5	55.6	9	2	Q12098_CAEV	Q12098	caprine art	773	5	55.6	10	2	Q6LB24_HUMAN	Q6LB24	homo sapien
701	5	55.6	9	2	Q12100_CAEV	Q12100	caprine art	774	5	55.6	10	2	Q6LC14_HUMAN	Q6LC14	homo sapien
702	5	55.6	9	2	Q12102_CAEV	Q12102	caprine art	775	5	55.6	10	2	Q6LEM6_HUMAN	Q6LEM6	homo sapien
703	5	55.6	9	2	Q12104_CAEV	Q12104	caprine art	776	5	55.6	10	2	Q6MZE6_HUMAN	Q6MZE6	homo sapien
704	5	55.6	9	2	Q64972_AVEVR	Q64972	avian rous-	777	5	55.6	10	2	Q70LT3_HUMAN	Q70LT3	homo sapien
705	5	55.6	9	2	Q70140_9HIV1	Q70140	human immun	778	5	55.6	10	2	Q7KZ15_HUMAN	Q7KZ15	homo sapien
706	5	55.6	9	2	Q8S599_MLYMO	Q8S599	moloney mur	779	5	55.6	10	2	Q7Z5A2_HUMAN	Q7Z5A2	homo sapien
707	5	55.6	9	2	Q8UTD7_9HIV1	Q8UTD7	human immun	780	5	55.6	10	2	Q8N6B1_HUMAN	Q8N6B1	homo sapien
708	5	55.6	10	2	Q8AEW8_9HIV1	Q8AEW8	human immun	781	5	55.6	10	2	Q8NEY9_HUMAN	Q8NEY9	homo sapien
709	5	55.6	10	1	AKHX_LOCFM1	P81626	locusta mig	782	5	55.6	10	2	Q8WXB5_HUMAN	Q8WXB5	homo sapien
710	5	55.6	10	1	ANGT1_BOTJA	Q10581	bothrops ja	783	5	55.6	10	2	Q8WXC4_HUMAN	Q8WXC4	homo sapien
711	5	55.6	10	1	ANGT1_BOVIN	P01017	bos taurus	784	5	55.6	10	2	Q9UCU6_HUMAN	Q9UCU6	homo sapien
712	5	55.6	10	1	ANGT_CHICK	P67885	gallus gall	785	5	55.6	10	2	Q9UD88_HUMAN	Q9UD88	homo sapien
713	5	55.6	10	1	ANGT_COTJA	P67886	coturnix co	786	5	55.6	10	2	Q9UNF2_HUMAN	Q9UNF2	homo sapien
714	5	55.6	10	1	ANOP_ANOSM	P0C005	anoplus sa	787	5	55.6	10	2	Q53TW5_HUMAN	Q53TW5	homo sapien
715	5	55.6	10	1	APE_CAFGI	P80474	capnocytoph	788	5	55.6	10	2	Q5S4Q0_HUMAN	Q5S4Q0	homo sapien
716	5	55.6	10	1	BPP2_BOTJA	P01022	bothrops ja	789	5	55.6	10	2	Q86XP4_HUMAN	Q86XP4	homo sapien
717	5	55.6	10	1	BRK_ONCV	Q9PRZ1	oncorhynch	790	5	55.6	10	2	Q4ZFF9_HUMAN	Q4ZFF9	homo sapien
718	5	55.6	10	1	CATB_SHEEP	P83205	ovis aries	791	5	55.6	10	2	Q5JPL1_HUMAN	Q5JPL1	homo sapien
719	5	55.6	10	1	COX81_CANFA	P61304	canis famil	792	5	55.6	10	2	Q3S013_MELJA	Q3S013	meloiodogone
720	5	55.6	10	1	COX81_RABIT	P80336	oryctolagus	793	5	55.6	10	2	Q5BRP3_SCHJA	Q5BRP3	schistosoma
721	5	55.6	10	1	COX82_CANFA	P61905	canis famil	794	5	55.6	10	2	Q5C1M9_SCHJA	Q5C1M9	schistosoma
722	5	55.6	10	1	FAR2_PENMO	P83317	penaeus mon	795	5	55.6	10	2	Q5C4A5_SCHJA	Q5C4A5	schistosoma
723	5	55.6	10	1	FAR5_MACRS	P83278	macrobrachi	796	5	55.6	10	2	P82383_DROME	P82383	drosophila
724	5	55.6	10	1	FAR6_PANRE	P82660	panagrellus	797	5	55.6	10	2	Q7MAC2_ECHMA	Q7MAC2	echinometra
725	5	55.6	10	1	FAR7_MACRS	P83280	macrobrachi	798	5	55.6	10	2	Q86D30_TRYCR	Q86D30	trypanosoma
726	5	55.6	10	1	FARC_CALVO	P41867	calliphora	799	5	55.6	10	2	Q8WPL6_9UROK	Q8WPL6	oikopleura
727	5	55.6	10	1	FARP_LOCFM1	P84306	locusta mig	800	5	55.6	10	2	Q9TWU1_FUSFE	Q9TWU1	fusinus fer
728	5	55.6	10	1	FARP_MANSE	P18523	manduca sex	801	5	55.6	10	2	Q7TWX1_DERFA	Q7TWX1	dermatophag
729	5	55.6	10	1	FARP_MYTED	P42560	mytilus edu	802	5	55.6	10	2	Q7RBG5_PLAYO	Q7RBG5	plasmodium
730	5	55.6	10	1	FARP_SCHGR	P84307	schistocerc	803	5	55.6	10	2	Q7RDS6_PLAYO	Q7RDS6	plasmodium
731	5	55.6	10	1	FIBB_CERSI	P14337	ceratotheri	804	5	55.6	10	2	Q7RKS4_PLAYO	Q7RKS4	plasmodium
732	5	55.6	10	1	GS09_BACSU	P80243	bacillus su	805	5	55.6	10	2	Q4XE58_PLACH	Q4XE58	plasmodium
733	5	55.6	10	1	LCMS_LEUMA	P21144	leucophaea	806	5	55.6	10	2	Q4XZ43_PLACH	Q4XZ43	plasmodium
734	5	55.6	10	1	LPK2_LOCFM1	P41488	locusta mig	807	5	55.6	10	2	Q4Z638_PLABE	Q4Z638	plasmodium
735	5	55.6	10	1	LSK2_LEUMA	P67802	leucophaea	808	5	55.6	10	2	Q5OS14_DIDMR	Q5OS14	didelphis m
736	5	55.6	10	1	LSK2_PERAM	P67803	periplaneta	809	5	55.6	10	2	Q6IG13_BOVIN	Q6IG13	bos taurus
737	5	55.6	10	1	MALE_KURPN	Q05564	klebsiella	810	5	55.6	10	2	Q7M2U1_BOVIN	Q7M2U1	bos taurus
738	5	55.6	10	1	MP2_MICOC	P81533	microplitis	811	5	55.6	10	2	Q8SPN8_MACMU	Q8SPN8	macaca mula
739	5	55.6	10	1	NEMS_SARBU	P61850	sarcophaga	812	5	55.6	10	2	Q9SH99_PACAN	Q9SH99	papio anubi
740	5	55.6	10	1	PA66D_SHEEP	P83496	ovis aries	813	5	55.6	10	2	Q9SHF4_PAPAN	Q9SHF4	papio anubi
741	5	55.6	10	1	PORB_METTM	P80901	methanobact	814	5	55.6	10	2	Q9SHF5_PAPAN	Q9SHF5	papio anubi
742	5	55.6	10	1	PSBF_CAPAN	Q03367	capricum an	815	5	55.6	10	2	Q9SHF6_PAPAN	Q9SHF6	papio anubi
743	5	55.6	10	1	PVK_LOCFM1	P83382	locusta mig	816	5	55.6	10	2	Q9SHF7_PAPAN	Q9SHF7	papio anubi
744	5	55.6	10	1	PVK_PHYMO	P84442	phymateus m	817	5	55.6	10	2	Q9SHF8_PAPAN	Q9SHF8	papio anubi
745	5	55.6	10	1	Q2OB_COMTE	P80465	comanotus t	818	5	55.6	10	2	Q9SHF9_PAPAN	Q9SHF9	papio anubi
746	5	55.6	10	1	RL16_ACHLA	P29221	acholeplasm	819	5	55.6	10	2	Q9SHG0_PAPAN	Q9SHG0	papio anubi
747	5	55.6	10	1	RNAM_PLESA	P84528	pleurotus s	820	5	55.6	10	2	Q9SM70_TRIVU	Q9SM70	trichosurus
748	5	55.6	10	1	RT02_BOVIN	P82923	bos taurus	821	5	55.6	10	2	Q9GKI4_MACAR	Q9GKI4	macaca arct
749	5	55.6	10	1	TKL2_LOCFM1	P16224	locusta mig	822	5	55.6	10	2	Q9GKI5_PANTR	Q9GKI5	pan troglod
750	5	55.6	10	1	TKL3_LOCFM1	P30249	locusta mig	823	5	55.6	10	2	Q9TRV4_HORSE	Q9TRV4	e equus cab
751	5	55.6	10	1	TKL4_LOCFM1	P30250	locusta mig	824	5	55.6	10	2	Q9TRV3_PIG	Q9TRV3	sus scrofa
752	5	55.6	10	1	TKN_PHYBI	P08610	phyllomedus	825	5	55.6	10	2	Q6JDL6_CANFA	Q6JDL6	canis famil
753	5	55.6	10	1	TKUI_UREUN	P40751	urechig uni	826	5	55.6	10	2	Q9XS84_HORSE	Q9XS84	equus cabal
754	5	55.6	10	1	TKU2_UREUN	P40752	urechig uni	827	5	55.6	10	2	Q7M2N0_BOVIN	Q7M2N0	bos taurus
755	5	55.6	10	1	TP1S_NICPL	P19118	nicotiana p	828	5	55.6	10	2	Q7M3E8_PIG	Q7M3E8	sus scrofa
756	5	55.6	10	1	TRP5_LEUMA	P81737	leucophaea	829	5	55.6	10	2	Q38217_9CAUD	Q38217	lactococcus
757	5	55.6	10	1	TRP6_LEUMA	P81738	leucophaea	830	5	55.6	10	2	Q5B4Y2_9MYRT	Q5B4Y2	ginoria ame
758	5	55.6	10	1	TRP7_LEUMA	P81739	leucophaea	831	5	55.6	10	2	Q518T2_CAMSI	Q518T2	camellia si
759	5	55.6	10	1	TRP8_LEUMA	P81740	leucophaea	832	5	55.6	10	2	Q518T5_9ERIC	Q518T5	camellia te
760	5	55.6	10	1	TRP9_LEUMA	P81741	leucophaea	833	5	55.6	10	2	Q518T6_9ERIC	Q518T6	camellia fu
761	5	55.6	10	1	UPA5_HUMAN	P30091	homo sapien	834	5	55.6	10	2			

835	5	55.6	10	2	Q518T7_CAMSI	Q518T7	camellia ei	908	5	55.6	10	2	P82588_STRPY	P82588	streptococc
836	5	55.6	10	2	P82434_TOBAC	P82434	nicotiana t	909	5	55.6	10	2	P83066_BACCE	P83066	bacillus ce
837	5	55.6	10	2	P82443_TOBAC	P82443	nicotiana t	910	5	55.6	10	2	P83067_BACCE	P83067	bacillus ce
838	5	55.6	10	2	P82937_HORVU	P82937	hordeum vul	911	5	55.6	10	2	P83160_ANASU	P83160	anabaena sp
839	5	55.6	10	2	P82938_HORVU	P82938	hordeum vul	912	5	55.6	10	2	P7M0J3_MARPU	P7M0J3	marichromat
840	5	55.6	10	2	Q5Y9E0_9BRVO	Q5Y9E0	scorpidium	913	5	55.6	10	2	P7M0M6_DESMA	P7M0M6	desulfovibr
841	5	55.6	10	2	Q5Y9F4_9BRVO	Q5Y9F4	calliergon	914	5	55.6	10	2	P7M0N4_DESMA	P7M0N4	serratia ma
842	5	55.6	10	2	Q6JVP0_WOLBI	Q6JVP0	wollastonia	915	5	55.6	10	2	Q847B5_BACAM	Q847B5	bacillus am
843	5	55.6	10	2	Q6JVP3_9ASTR	Q6JVP3	otopappus e	916	5	55.6	10	2	Q8KHN9_CLOBO	Q8KHN9	clostridium
844	5	55.6	10	2	Q6JVP6_9ASTR	Q6JVP6	elaphandra	917	5	55.6	10	2	Q8RUF1_PSEFL	Q8RUF1	pseudomonas
845	5	55.6	10	2	Q6JVP8_9ASTR	Q6JVP8	dimerostemm	918	5	55.6	10	2	Q931X4_VIBCH	Q931X4	vibrio chol
846	5	55.6	10	2	Q6JVP0_9ASTR	Q6JVP0	dimerostemm	919	5	55.6	10	2	Q93UU2_ECOS7	Q93UU2	escherichia
847	5	55.6	10	2	Q6JVP3_9ASTR	Q6JVP3	baltimora r	920	5	55.6	10	2	Q9JNC9_STRPY	Q9JNC9	streptococc
848	5	55.6	10	2	Q6JVP5_9ASTR	Q6JVP5	angelphytum	921	5	55.6	10	2	Q9RSN1_CLOBO	Q9RSN1	clostridium
849	5	55.6	10	2	Q6JVP7_9ASTR	Q6JVP7	angelphytum	922	5	55.6	10	2	Q9RSN3_CLOBO	Q9RSN3	clostridium
850	5	55.6	10	2	Q6KCG9_EUGCL	Q6KCG9	eucalyptus	923	5	55.6	10	2	Q9R7J9_HELPY	Q9R7J9	helicobacte
851	5	55.6	10	2	Q70Y78_LAMI	Q70Y78	plectranthu	924	5	55.6	10	2	Q9S3J6_ECOLI	Q9S3J6	escherichia
852	5	55.6	10	2	Q7M1F6_9POAL	Q7M1F6	haynaldia v	925	5	55.6	10	2	Q9X533_ECOLI	Q9X533	escherichia
853	5	55.6	10	2	Q7M1F7_9POAL	Q7M1F7	haynaldia v	926	5	55.6	10	2	Q9X534_9ENTR	Q9X534	leclercia a
854	5	55.6	10	2	Q7M1I1_PHAVU	Q7M1I1	phaseolus v	927	5	55.6	10	2	Q54217_STRAP	Q54217	staphylococ
855	5	55.6	10	2	Q7M1I6_TRIKI	Q7M1I6	trichobanth	928	5	55.6	10	2	Q71V02_PSEAB	Q71V02	pseudomonas
856	5	55.6	10	2	Q7M1V8_NICPL	Q7M1V8	nicotiana p	929	5	55.6	10	2	Q6JL97_NEIGO	Q6JL97	neisseria g
857	5	55.6	10	2	Q7M278_TRITU	Q7M278	triticum tu	930	5	55.6	10	2	P74843_STRTH	P74843	streptomyce
858	5	55.6	10	2	Q7M282_ORXSA	Q7M282	oryza kativ	931	5	55.6	10	2	Q8VN85_HELPY	Q8VN85	helicobacte
859	5	55.6	10	2	Q7M2G1_VICFA	Q7M2G1	vicia faba	932	5	55.6	10	2	Q48469_KLEPN	Q48469	klebsiella
860	5	55.6	10	2	Q7M2G2_VICFA	Q7M2G2	vicia faba	933	5	55.6	10	2	Q79AV7_KLEPN	Q79AV7	klebsiella
861	5	55.6	10	2	Q85AZ9_9BRVO	Q85AZ9	pyrrhobryum	934	5	55.6	10	2	Q5D4R6_9BURK	Q5D4R6	alcaligenes
862	5	55.6	10	2	Q85BV5_EUCGR	Q85BV5	eucalyptus	935	5	55.6	10	2	Q7WUG1_PSEFL	Q7WUG1	pseudomonas
863	5	55.6	10	2	Q85BV6_EUCGR	Q85BV6	eucalyptus	936	5	55.6	10	2	Q08622_RAT	Q08622	rattus norv
864	5	55.6	10	2	Q85BV7_EUCGR	Q85BV7	eucalyptus	937	5	55.6	10	2	Q63389_RAT	Q63389	rattus norv
865	5	55.6	10	2	Q85V65_EUCGR	Q85V65	eucalyptus	938	5	55.6	10	2	Q68SM8_CHAPN	Q68SM8	chaetodipus
866	5	55.6	10	2	Q85V66_EUCGR	Q85V66	eucalyptus	939	5	55.6	10	2	Q68SM9_THOMO	Q68SM9	thomomys mo
867	5	55.6	10	2	Q85V67_EUCGR	Q85V67	eucalyptus	940	5	55.6	10	2	Q68SN0_CRAGY	Q68SN0	cratogeomys
868	5	55.6	10	2	Q8GZC8_HORVU	Q8GZC8	hordeum vul	941	5	55.6	10	2	Q68SN1_CRAGY	Q68SN1	cratogeomys
869	5	55.6	10	2	Q8HUB4_9BRVO	Q8HUB4	anomobkium	942	5	55.6	10	2	Q6LD58_9MURI	Q6LD58	mus sp. alp
870	5	55.6	10	2	Q8MAZ9_9ASTE	Q8MAZ9	dicrangetyl	943	5	55.6	10	2	Q6S510_MOUSE	Q6S510	mus musculu
871	5	55.6	10	2	Q8MBB7_9ASTE	Q8MBB7	merremia ae	944	5	55.6	10	2	Q80WD4_MOUSE	Q80WD4	mus musculu
872	5	55.6	10	2	Q8SAC2_9BRVO	Q8SAC2	amblystegiu	945	5	55.6	10	2	Q80WD9_9MURI	Q80WD9	rattus sp.
873	5	55.6	10	2	Q947R7_SOLTU	Q947R7	solanum tub	946	5	55.6	10	2	Q80Z98_RAT	Q80Z98	rattus norv
874	5	55.6	10	2	Q94I19_MAIZE	Q94I19	zea mays (m	947	5	55.6	10	2	Q8BHN2_MOUSE	Q8BHN2	mus musculu
875	5	55.6	10	2	Q9THM6_9MYRT	Q9THM6	leptospermu	948	5	55.6	10	2	Q8CIN5_CAVPO	Q8CIN5	cavia porce
876	5	55.6	10	2	Q9THM7_9MYRT	Q9THM7	leptospermu	949	5	55.6	10	2	Q8CJ31_MOUSE	Q8CJ31	mus musculu
877	5	55.6	10	2	Q9TKB0_9MYRT	Q9TKB0	neofabricia	950	5	55.6	10	2	Q91WZ3_9MURI	Q91WZ3	rattus sp.
878	5	55.6	10	2	Q9TKB1_9MYRT	Q9TKB1	neofabricia	951	5	55.6	10	2	Q9QVE5_9MURI	Q9QVE5	mus sp. pro
879	5	55.6	10	2	Q9TKB2_9MYRT	Q9TKB2	leptospermu	952	5	55.6	10	2	Q9QVE6_9MURI	Q9QVE6	mus sp. pro
880	5	55.6	10	2	Q9TKB3_9MYRT	Q9TKB3	leptospermu	953	5	55.6	10	2	Q9QVE7_9MURI	Q9QVE7	mus sp. pro
881	5	55.6	10	2	Q9TKB4_9MYRT	Q9TKB4	leptospermu	954	5	55.6	10	2	Q9QVE8_9MURI	Q9QVE8	mus sp. pro
882	5	55.6	10	2	Q9TKB5_9MYRT	Q9TKB5	leptospermu	955	5	55.6	10	2	Q9QVF0_9MURI	Q9QVF0	mus sp. pro
883	5	55.6	10	2	Q9TKB6_9MYRT	Q9TKB6	leptospermu	956	5	55.6	10	2	Q9QVF1_9MURI	Q9QVF1	mus sp. pro
884	5	55.6	10	2	Q9TKB7_KUNPU	Q9TKB7	kunzea pulc	957	5	55.6	10	2	Q9QVF7_9MURI	Q9QVF7	rattus sp.
885	5	55.6	10	2	Q9TKB8_KUNBA	Q9TKB8	kunzea eric	958	5	55.6	10	2	Q9T2P3_9MURI	Q9T2P3	rattus sp.
886	5	55.6	10	2	Q9TKB9_KUNBA	Q9TKB9	kunzea baxt	959	5	55.6	10	2	Q922V3_RAT	Q922V3	rattus norv
887	5	55.6	10	2	Q9TKF0_KUNAM	Q9TKF0	kunzea ambi	960	5	55.6	10	2	Q53VQ3_MOUSE	Q53VQ3	mus musculu
888	5	55.6	10	2	Q9TKF1_9MYRT	Q9TKF1	homalopherm	961	5	55.6	10	2	Q53VQ7_MOUSE	Q53VQ7	mus musculu
889	5	55.6	10	2	Q9TKF3_9MYRT	Q9TKF3	asteromyrtu	962	5	55.6	10	2	Q53VR1_MOUSE	Q53VR1	mus musculu
890	5	55.6	10	2	Q9TKF4_9MYRT	Q9TKF4	angasomyrtu	963	5	55.6	10	2	Q7TSC5_MOUSE	Q7TSC5	mus musculu
891	5	55.6	10	2	Q9TKF5_9MYRT	Q9TKF5	agonis spat	964	5	55.6	10	2	Q5TLE7_MOUSE	Q5TLE7	mus musculu
892	5	55.6	10	2	Q9TKF6_9MYRT	Q9TKF6	agonis obtu	965	5	55.6	10	2	Q6LBT3_MOUSE	Q6LBT3	mus musculu
893	5	55.6	10	2	Q9TKF7_9MYRT	Q9TKF7	agonis gran	966	5	55.6	10	2	Q80WD3_MUSSP	Q80WD3	mus spretus
894	5	55.6	10	2	Q9TKF8_9MYRT	Q9TKF8	tristahiops	967	5	55.6	10	2	Q90346_9FLAV	Q90346	gb virus c/
895	5	55.6	10	2	Q9TKF9_MELVI	Q9TKF9	melaleuca v	968	5	55.6	10	2	O90347_9FLAV	O90347	gb virus c/
896	5	55.6	10	2	Q9TKG0_9MYRT	Q9TKG0	lophostemon	969	5	55.6	10	2	O90348_9FLAV	O90348	gb virus c/
897	5	55.6	10	2	Q9TKG2_CALPO	Q9TKG2	callistemon	970	5	55.6	10	2	O90349_9FLAV	O90349	gb virus c/
898	5	55.6	10	2	Q9XMB4_AEGTA	Q9XMB4	aegilops ta	971	5	55.6	10	2	O64971_9BROM	O64971	alfalfa mos
899	5	55.6	10	2	Q56ZK9_ARATH	Q56ZK9	arabidopsis	972	5	55.6	10	2	Q66M69_9VIRU	Q66M69	potato viru
900	5	55.6	10	2	Q5SBS2_9POAL	Q5SBS2	thamnochoit	973	5	55.6	10	2	Q66M69_9VIRU	Q66M69	potato viru
901	5	55.6	10	2	P82132_SPIOL	P82132	spinacia ol	974	5	55.6	10	2	Q69347_HVVI	Q69347	human herpe
902	5	55.6	10	2	P82133_SPIOL	P82133	spinacia ol	975	5	55.6	10	2	Q70GQ6_PRYKA	Q70GQ6	pseudorabie
903	5	55.6	10	2	P82136_SPIOL	P82136	spinacia ol	976	5	55.6	10	2	Q76V79_POLY	Q76V79	polyomaviru
904	5	55.6	10	2	Q4QWV8_9MARC	Q4QWV8	jensenia di	977	5	55.6	10	2	Q53X10_POVBK	Q53X10	polyomaviru
905	5	55.6	10	2	Q44693_BACNM	Q44693	bacillus am	978	5	55.6	10	2	Q53X11_POVBK	Q53X11	polyomaviru
906	5	55.6	10	2	Q50032_MYCLE	Q50032	mycobacteri	979	5	55.6	10	2	Q53X33_POVBK	Q53X33	polyomaviru
907	5	55.6	10	2	Q52B37_RHILU	Q52B37	rhizobium l	980	5	55.6	10	2	Q82625_IBDV	Q82625	avian infec
					Q5D4Q3_9RHOO	Q5D4Q3	azospira or								

```

981 5 55.6 10 2 042355_BRARE
982 5 55.6 10 2 073589_CHICK
983 5 55.6 10 2 073594_CHICK
984 5 55.6 10 2 079885_9SAUR
985 5 55.6 10 2 079888_BASPL
986 5 55.6 10 2 079891_CROCO
987 5 55.6 10 2 079894_GAMWI
988 5 55.6 10 2 079897_HOPSP
989 5 55.6 10 2 079900_9SAUR
990 5 55.6 10 2 079903_OPLCU
991 5 55.6 10 2 079906_Phrynosoma
992 5 55.6 10 2 079909_sauronulus
993 5 55.6 10 2 079912_CHAFI
994 5 55.6 10 2 079915_LEIBE
995 5 55.6 10 2 079924_9SAUR
996 5 55.6 10 2 P92576_BIPBI
997 5 55.6 10 2 P92616_Aspidoceli
998 5 55.6 10 2 P92648_9SAUR
999 5 55.6 10 2 P92758_teratoscinc
1000 5 55.6 10 2 P92762_uromastyx a

```

## ALIGNMENTS

```

RESULT 1
FAR3_HIRME STANDARD; PRT; 4 AA.
AC P42562;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE FMRFamide-like neuropeptide YLRF-amide.
OS Hirudo medicinalis (Medicinal leech).
OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
OC Arhynchobdellida; Hirudiniformes; Hirudinidae; Hirudo.
OX NCBI_TaxID=6421;
RN [1]
RP PROTEIN SEQUENCE.
RX MEDLINE=92195954; PubMed=1686933; DOI=10.1016/0196-9781(91)90035-N;
RA Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
RT "Identification of RFamide neuropeptides in the medicinal leech.";
RL Peptides 12:897-908(1991).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide) family.

```

```

-----
This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
-----
KW Amidation; Direct protein sequencing; Neuropeptide.
FT MOD RES 4 4 Phenylalanine amide.
SQ SEQUENCE 4 AA; 598 MW; 69D4073B30000000 CRC64;

```

```

Query Match 55.6%; Score 5; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 R 1
Db 3 R 3

```

```

RESULT 2
FAR4_HIRME STANDARD; PRT; 4 AA.
ID FAR4_HIRME
AC P42563;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)

```

```

DE FMRFamide-like neuropeptide YLRF-amide.
OS Hirudo medicinalis (Medicinal leech).
OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
OC Arhynchobdellida; Hirudiniformes; Hirudinidae; Hirudo.
OX NCBI_TaxID=6421;
RN [1]
RP PROTEIN SEQUENCE.
RX MEDLINE=92195954; PubMed=1686933; DOI=10.1016/0196-9781(91)90035-N;
RA Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
RT "Identification of RFamide neuropeptides in the medicinal leech.";
RL Peptides 12:897-908(1991).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide) family.
-----
This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
-----
KW Amidation; Direct protein sequencing; Neuropeptide.
FT MOD RES 4 4 Phenylalanine amide.
SQ SEQUENCE 4 AA; 616 MW; 69D4068B30000000 CRC64;

```

```

Query Match 55.6%; Score 5; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 R 1
Db 3 R 3

```

```

RESULT 3
FLRF_HELTI STANDARD; PRT; 4 AA.
ID FLRF_HELTI
AC P69138; P42561;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE FLRFamide.
OS Helisoma trivolvis (Snail).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;
OC Lymnaeidae; Planorbidae; Helisoma.
OX NCBI_TaxID=27815;
RN [1]
RP PROTEIN SEQUENCE.
RX TISSUE=Kidney.
RX MEDLINE=94286417; PubMed=79124228; DOI=10.1016/0196-9781(94)90166-X;
RA Madrid K.P., Price D.A., Greenberg M.J., Khan H.R., Saleuddin A.S.M.;
RT "FMRFamide-related peptides from the kidney of the snail, Helisoma
trivolvis.";
RL Peptides 15:31-36(1994).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide) family.

```

```

-----
This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
-----
KW Amidation; Direct protein sequencing; Neuropeptide.
FT MOD RES 4 4 Phenylalanine amide.
SQ SEQUENCE 4 AA; 582 MW; 69D40729A0000000 CRC64;

```

```

Query Match 55.6%; Score 5; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 R 1

```

```

Db      | 3 R 3
|
RESULT 4
FLRF_HIRME
ID FLRF_HIRME STANDARD; PRT; 4 AA.
AC P69137; P42561;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE FLRFamide.
OS Hirudo medicinalis (Medicinal leech).
OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
OC Arhynchobdellida; Hirudiniformes; Hirudinidae; Hirudo.
OX NCBI_TaxID=6421;
RN [1]
RP PROTEIN SEQUENCE.
RX MEDLINE=92195954; PubMed=1686933; DOI=10.1016/0196-9781(91)90035-N;
RA Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
RT "Identification of RFamide neuropeptides in the medicinal leech.";
RL Peptides 12:897-908(1991).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)
CC family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC Amidation; Direct protein sequencing; Neuropeptide.
KW MOD_RES 4 4 Phenylalanine amide.
FT MOD_RES 4 4
SQ SEQUENCE 4 AA; 582 MW; 69D40729A0000000 CRC64;

Query Match 55.6%; Score 5; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1
|
Db 3 R 3

RESULT 5
FLRN_ATEL
ID FLRN_ATEL STANDARD; PRT; 4 AA.
AC P58707;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Antho-RNamide.
OS Anthopleura elegantissima (Sea anemone).
OC Eukaryota; Metazoa; Chnidaria; Anthozoa; Zoantharia; Actiniaria;
OC Nynanthaeae; Actiniidae; Anthopleura.
OX NCBI_TaxID=6110;
RN [1]
RP PROTEIN SEQUENCE, AND MASS SPECTROMETRY.
RX MEDLINE=90319122; PubMed=1973541;
RA Grimmelikhuijzen C.J.P., Rinehart K.L. Jr., Jacob E., Graff D.,
RA Reinscheid R.K., Nothacker H.-P., Staley A.L.;
RT "Isolation of L-3-phenylacetyl-Leu-Arg-Asn-NH2 (Antho-RNamide), a sea
RT anemone neuropeptide containing an unusual amino-terminal blocking
RT group.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:5410-5414(1990).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Neuron specific.
CC -!- MASS SPECTROMETRY: MW=549.3; METHOD=PAB; RANGE=1-4; NOTE=Ref.1.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC Amidation; Direct protein sequencing; Neuropeptide.
KW MOD_RES 4 4 Phenylalanine amide.
FT MOD_RES 4 4
SQ SEQUENCE 4 AA; 582 MW; 69D40729A0000000 CRC64;

Query Match 55.6%; Score 5; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1
|
Db 3 R 3

RESULT 6
FMRF_HELTI
ID FMRF_HELTI STANDARD; PRT; 4 AA.
AC P69148; P01162;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE FMRFamide.
OS Helisoma trivolvis (Snail).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;
OC Lymnaeidae; Planorbidae; Helisoma.
OX NCBI_TaxID=27815;
RN [1]
RP PROTEIN SEQUENCE.
RX TISSUE=Kidney;
RX MEDLINE=94286417; PubMed=7912428; DOI=10.1016/0196-9781(94)90166-X;
RA Madril K.P., Price D.A., Greenberg M.J., Khan H.R., Saleuddin A.S.M.;
RT "FMRFamide-related peptides from the kidney of the snail, Helisoma
RT trivolvis.";
RL Peptides 15:31-36(1994).
CC -!- FUNCTION: Myoactive; cardioexcitatory substance. Pharmacological
CC activities include augmentation, induction, and regularization of
CC cardiac contraction.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)
CC family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC Amidation; Direct protein sequencing; Neuropeptide.
KW MOD_RES 4 4 Phenylalanine amide.
FT MOD_RES 4 4
SQ SEQUENCE 4 AA; 600 MW; 69D40699A0000000 CRC64;

Query Match 55.6%; Score 5; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1
|
Db 3 R 3

RESULT 7
FMRP_HIRME
ID FMRP_HIRME STANDARD; PRT; 4 AA.
AC P69147; P01162;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE FMRFamide.
OS Hirudo medicinalis (Medicinal leech).

```

CC use as long as its content is in no way modified and this statement is not removed.

DR PIR; A35779; A35779.  
 KW Amidation; Direct protein sequencing; Neuropeptide.  
 FT MOD\_RES 1 1 3-phenyllactic acid.  
 FT MOD\_RES 4 4 Asparagine amide.  
 SQ SEQUENCE 4 AA; 549 MW; 64540729A0000000 CRC64;

Query Match 55.6%; Score 5; DB 1; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1  
 |  
 Db 3 R 3

RESULT 6  
 FMRF\_HELTI  
 ID FMRF\_HELTI STANDARD; PRT; 4 AA.  
 AC P69148; P01162;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 10-MAY-2005 (Rel. 47, Last annotation update)  
 DE FMRFamide.

OS Helisoma trivolvis (Snail).  
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;  
 OC Lymnaeidae; Planorbidae; Helisoma.  
 OX NCBI\_TaxID=27815;  
 RN [1]  
 RP PROTEIN SEQUENCE.

RX TISSUE=Kidney;  
 RX MEDLINE=94286417; PubMed=7912428; DOI=10.1016/0196-9781(94)90166-X;  
 RA Madril K.P., Price D.A., Greenberg M.J., Khan H.R., Saleuddin A.S.M.;  
 RT "FMRFamide-related peptides from the kidney of the snail, Helisoma  
 RT trivolvis";  
 RL Peptides 15:31-36(1994).

CC -!- FUNCTION: Myoactive; cardioexcitatory substance. Pharmacological  
 CC activities include augmentation, induction, and regularization of  
 CC cardiac contraction.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)  
 CC family.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.

CC Amidation; Direct protein sequencing; Neuropeptide.  
 KW MOD\_RES 4 4 Phenylalanine amide.  
 FT MOD\_RES 4 4  
 SQ SEQUENCE 4 AA; 600 MW; 69D40699A0000000 CRC64;

Query Match 55.6%; Score 5; DB 1; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1  
 |  
 Db 3 R 3

RESULT 7  
 FMRP\_HIRME  
 ID FMRP\_HIRME STANDARD; PRT; 4 AA.  
 AC P69147; P01162;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 10-MAY-2005 (Rel. 47, Last annotation update)  
 DE FMRFamide.  
 OS Hirudo medicinalis (Medicinal leech).

CC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;  
CC Arynchobdellida; Hirudiniiformes; Hirudinidae; Hirudo.  
OX NCBI\_TaxID=6421;  
RN [1]  
RP PROTEIN SEQUENCE.  
RX MEDLINE=92195954; PubMed=1686933; DOI=10.1016/0196-9781(91)90035-N;  
RA Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;  
RT "Identification of Rfamade neuropeptides in the medicinal leech.";  
RL Peptides 12:897-908(1991).  
CC -!- FUNCTION: Myoactive; cardioexcitatory substance. Pharmacological  
CC activities include augmentation, induction, and regularization of  
CC cardiac contraction.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)  
CC family.  
CC  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC  
KW Amidation; Direct protein sequencing; Neuropeptide.  
FT MOD RES 4 4 Phenylalanine amide.  
SQ SEQUENCE 4 AA; 600 MW; 69D40699A0000000 CRC64;  
  
Query Match 55.6%; Score 5; DB 1; Length 4;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 R 1  
DB 3 R 3  
  
RESULT 8  
FMRP\_MACNI STANDARD; PRT; 4 AA.  
ID FMRP MACNI  
AC P69145; P01162;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 10-MAY-2005 (Rel. 47, Last annotation update)  
DE FMRFamide (peak C) (cardioexcitatory neuropeptide).  
OS Macrocallista nimbosa (Sun-ray clam).  
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Heteroconchia; Veneroidea;  
OC Veneroidea; Veneridae; Macrocallista.  
OX NCBI\_TaxID=6594;  
RN [1]  
RP PROTEIN SEQUENCE, AND SYNTHESIS.  
RC TISSUE=Cerebral pedal, and Visceral ganglion;  
RX MEDLINE=77215956; PubMed=877582;  
RA Price D.A., Greenberg M.J.;  
RT "Structure of a molluscan cardioexcitatory neuropeptide.";  
RL Science 197:670-671(1977).  
RN [2]  
RP PROTEIN SEQUENCE, AND CHARACTERIZATION.  
RC TISSUE=Ganglion;  
RX MEDLINE=78012038; PubMed=909875;  
RA Price D.A., Greenberg M.J.;  
RT "Purification and characterization of a cardioexcitatory neuropeptide  
from the central ganglia of a bivalve mollusc.";  
RL Prep. Biochem. 7:261-281(1977).  
CC -!- FUNCTION: Myoactive; cardioexcitatory substance. Pharmacological  
CC activities include augmentation, induction, and regularization of  
CC cardiac contraction.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)  
CC family.  
CC  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not

CC removed.  
CC  
CC PIR; A01426; ECKN.  
KW Amidation; Direct protein sequencing; Neuropeptide.  
FT MOD RES 4 4 Phenylalanine amide.  
SQ SEQUENCE 4 AA; 600 MW; 69D40699A0000000 CRC64;  
  
Query Match 55.6%; Score 5; DB 1; Length 4;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 R 1  
DB 3 R 3  
  
RESULT 9  
FMRP\_NERVI STANDARD; PRT; 4 AA.  
ID FMRP NERVI  
AC P69146; P01162;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 10-MAY-2005 (Rel. 47, Last annotation update)  
DE FMRFamide.  
OS Nereis virens (Sandworm).  
OC Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Aciculata;  
OC Phyllocorida; Nereididae; Nereis.  
OX NCBI\_TaxID=6353;  
RN [1]  
RP PROTEIN SEQUENCE.  
RX MEDLINE=90259866; PubMed=2342992; DOI=10.1016/0196-9781(90)90113-J;  
RA Krajniak K.G., Price D.A.;  
RT "Authentic FMRFamide is present in the polychaete Nereis virens.";  
RL Peptides 11:75-77(1990).  
CC -!- FUNCTION: Myoactive; cardioexcitatory substance. Pharmacological  
CC activities include augmentation, induction, and regularization of  
CC cardiac contraction.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)  
CC family.  
CC  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC  
CC PIR; A60418; A60418.  
KW Amidation; Direct protein sequencing; Neuropeptide.  
FT MOD RES 4 4 Phenylalanine amide.  
SQ SEQUENCE 4 AA; 600 MW; 69D40699A0000000 CRC64;  
  
Query Match 55.6%; Score 5; DB 1; Length 4;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 R 1  
DB 3 R 3  
  
RESULT 10  
FYRI\_ANCEL STANDARD; PRT; 4 AA.  
ID FYRI ANCEL  
AC P58706;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 01-FEB-2005 (Rel. 46, Last annotation update)  
DE Antho-Riamide I [Contains: Antho-Riamide II].  
OS Anthopleura elegantissima (Sea anemone).  
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actinaria;  
OC Nynanthae; Actinellidae; Anthopleura.  
OX NCBI\_TaxID=6110;

RN PROTEIN SEQUENCE.  
 RX MEDLINE=92270459; PubMed=1821096; DOI=10.1016/0196-9781(91)90190-Z;  
 RA Nothacker H.-P., Rinehart K.L. Jr., McFarlane I.D.,  
 RA Grimmelikhuijzen C.J.P.;  
 RT "Isolation of two novel neuropeptides from sea anemones: the unusual,  
 RT biologically active L-3-phenylalanyl-Tyr-Arg-Ile-NH2 and its des-  
 RT phenylalanyl fragment Tyr-Arg-Ile-NH2.";  
 RL Peptides 12:1165-1173(1991).  
 RN [2]  
 RP FUNCTION.  
 RX MEDLINE=93391436; PubMed=8397415;  
 RA McFarlane I.D., Hudman D., Nothacker H.-P., Grimmelikhuijzen C.J.P.;  
 RT "The expansion behaviour of sea anemones may be coordinated by two  
 RT inhibitory neuropeptides, Antho-Kamide and Antho-Riamide.";  
 RL Proc. R. Soc. Lond., B, Biol. Sci. 253:183-188(1993).  
 CC -!- FUNCTION: Inhibits spontaneous contractions in several muscle  
 CC groups. May be involved in the expansion phase of feeding  
 CC behaviour in sea anemones.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Neuron specific.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC -----  
 KW Amidation; Direct protein sequencing; Neuropeptide.  
 FT PEPTIDE 1 4 Antho-Riamide I.  
 FT PEPTIDE 2 4 Antho-Riamide II.  
 FT MOD RES 1 1 3-phenyllactic acid.  
 FT MOD RES 4 4 Isoleucine amide.  
 SQ SEQUENCE 4 AA; 598 MW; 60441B59A0000000 CRC64;  
 Query Match 55.6%; Score 5; DB 1; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 R 1  
 DB 3 R 3  
 RESULT 11  
 TUFT\_HUMAN  
 ID TUFT\_HUMAN STANDARD; PRT; 4 AA.  
 AC P01858;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Phagocytosis-stimulating peptide (Tuftsin).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
 OC Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP PROTEIN SEQUENCE.  
 RX MEDLINE=72187087; PubMed=4112769;  
 RA Nishioka K., Constantopoulos A., Satoh P.S., Najjar V.A.;  
 RT "The characteristics, isolation and synthesis of the phagocytosis  
 RT stimulating peptide tuftsin.";  
 RL Biochem. Biophys. Res. Commun. 47:172-179(1972).  
 RN [2]  
 RP IMMUNOGLOBULIN CLASS.  
 RX MEDLINE=68091045; PubMed=4169272;  
 RA Fidalgo B.V., Najjar V.A.;  
 RT "The physiological role of the lymphoid system. VI. The stimulatory  
 RT effect of leucophilic gamma globulin (leucokinin) on the phagocytic  
 RT activity of human polymorphonuclear leucocyte.";  
 RL Biochemistry 6:3386-3392(1967).  
 CC -!- MISCELLANEOUS: An IgG (called leucokinin) binds reversibly to the

CC cell membrane of neutrophils in the blood. Leucokininase on the  
 CC membrane releases the active peptide tuftsin from the gamma chain.  
 CC Tuftsin is essential for maximum stimulation of the phagocytic  
 CC activity of neutrophils.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC -----  
 DR PIR; A02147; A02147.  
 DR MIM; 191150; .  
 DR GO; GO:0003823; F:antigen binding; NAS.  
 DR GO; GO:0006909; P:phagocytosis; NAS.  
 KW Direct protein sequencing.  
 SQ SEQUENCE 4 AA; 501 MW; 74176321C0000000 CRC64;  
 Query Match 55.6%; Score 5; DB 1; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 R 1  
 DB 4 R 4  
 RESULT 12  
 FARP\_ARTTR  
 ID FARP\_ARTTR STANDARD; PRT; 5 AA.  
 AC P41853;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE FMRamide-like neuropeptide RYRP-amide.  
 OS Artiposthia triangulata (New Zealand flatworm).  
 OC Eukaryota; Metazoa; Platyhelminthes; Turbellaria; Seriata; Tricladida;  
 OC Terricola; Geoplanidae; Arthurdendyus.  
 OX NCBI\_TaxID=132421;  
 RN [1]  
 RP PROTEIN SEQUENCE, AND SYNTHESIS.  
 RX MEDLINE=94211927; PubMed=7909164; DOI=10.1016/0167-0115(94)90189-9;  
 RA Maule A.G., Shaw C., Halton D.W., Curry W.J., Thim L.;  
 RT "RYRamide: a turbellarian FMRamide-related peptide (FARP).";  
 RL Regul. Pept. 50:37-43(1994).  
 CC -!- SIMILARITY: Belongs to the FARP (FMRamide related peptide)  
 CC family.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC -----  
 KW Amidation; Direct protein sequencing; Neuropeptide.  
 FT MOD RES 5 5 phenylalanine amide.  
 SQ SEQUENCE 5 AA; 754 MW; 69D4004B44600000 CRC64;  
 Query Match 55.6%; Score 5; DB 1; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 R 1  
 DB 1 R 1  
 RESULT 13  
 FARP\_CHICK  
 ID FARP\_CHICK STANDARD; PRT; 5 AA.  
 AC P83308;  
 DT 05-JUL-2004 (Rel. 44, Created)  
 DT 05-JUL-2004 (Rel. 44, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE FMRFamide-like neuropeptide (LPLRF-amide).  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OX Gallus.  
CN NCBI\_TaxID=9031;  
RN [1]  
RP PROTEIN SEQUENCE, AND SYNTHESIS.  
RC TISSUE=Brain;  
RX PubMed=613771;  
RA Dockray G.J., Reeve J.R. Jr., Shively J., Gayton R.J., Barnard C.S.;  
RT "A novel active pentapeptide from chicken brain identified by  
RT antibodies to FMRFamide.";  
RL Nature 305:328-330(1983).  
CC -!- FUNCTION: May function as a neurotransmitter or modulator.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)  
CC family.  
CC -----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC -----  
CC GO: GO:0007218; P:neuropeptide signaling pathway; TAS.  
DR Amidation; Direct protein sequencing; Neuropeptide.  
KW MOD RES 5 5 Phenylalanine amide.  
FT SEQUENCE 5 AA; 645 MW; 69D4073767400000 CRC64;  
SQ  
  
Query Match 55.6%; Score 5; DB 1; Length 5;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 R 1  
Db 4 R 4  
  
RESULT 14  
PCT\_CARMA  
ID PCT\_CARMA STANDARD; PRT; 5 AA.  
AC P67857; P01373;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Proctolin.  
OS Carcinus maenas (Common shore crab) (Green crab).  
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;  
OC Eubrachyura; Portunoidae; Portunidae; Carcinus.  
OX NCBI\_TaxID=6759;  
RN [1]  
RP PROTEIN SEQUENCE.  
RX MEDLINE=86232789; PubMed=2872661; DOI=10.1016/0196-9781(86)90063-X;  
RA Stangier J., Dircksen H., Keller R.;  
RT "Identification and immunocytochemical localization of proctolin in  
RT pericardial organs of the shore crab, Carcinus maenas.";  
RL Peptides 7:67-72(1986).  
CC -!- FUNCTION: Stimulates cardiac output and hindgut motility,  
CC modulates visceral and skeletal muscle in many arthropods.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Found in the crab pericardial organs.  
CC -----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC -----  
CC Direct protein sequencing; Neuropeptide.  
KW SEQUENCE 5 AA; 649 MW; 71B7673B44600000 CRC64;  
SQ

Query Match 55.6%; Score 5; DB 1; Length 5;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 R 1  
Db 1 R 1  
  
RESULT 15  
PCT\_LIMPO  
ID PCT\_LIMPO STANDARD; PRT; 5 AA.  
AC P67858; P01373;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Proctolin.  
OS Limulus polyphemus (Atlantic horseshoe crab).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Merostomata; Xiphosura;  
OC Limulidae; Limulus.  
OX NCBI\_TaxID=6850;  
RN [1]  
RP PROTEIN SEQUENCE.  
RX MEDLINE=90287800; PubMed=2356151; DOI=10.1016/0196-9781(90)90072-D;  
RA Groome J.R., Tillinghast E.K., Townley M.A., Vetrovs A.,  
RA Watson W.H. III, Hunt D.F., Griffin P.R., Alexander J.E.,  
RA Shabanowitz J.;  
RT "Identification of proctolin in the central nervous system of the  
RT horseshoe crab, Limulus polyphemus.";  
RL Peptides 11:205-211(1990).  
CC -!- FUNCTION: Stimulates cardiac output and hindgut motility,  
CC modulates visceral and skeletal muscle in many arthropods.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Found in the crab pericardial organs.  
CC -----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC -----  
CC PIR; A60411; A60411.  
DR Direct protein sequencing; Neuropeptide.  
KW SEQUENCE 5 AA; 649 MW; 71B7673B44600000 CRC64;  
SQ  
  
Query Match 55.6%; Score 5; DB 1; Length 5;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 R 1  
Db 1 R 1  
  
RESULT 16  
PCT\_PERAM  
ID PCT\_PERAM STANDARD; PRT; 5 AA.  
AC P67859; P01373;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Proctolin.  
OS Periplaneta americana (American cockroach).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blattodea;  
OC Blattidae; Blattellinae; Periplaneta.  
OX NCBI\_TaxID=6976;  
RN [1]  
RP PROTEIN SEQUENCE.  
RX MEDLINE=76074708; PubMed=576; DOI=10.1016/0024-3205(75)90134-4;  
RA Starratt A.N., Brown B.E.;  
RT "Structure of the pentapeptide proctolin, a proposed neurotransmitter



RT in insects";  
 RL Life Sci. 17:1253-1256 (1975).  
 RN [2]  
 RP BIOLOGICAL SOURCE.  
 RX MEDLINE=81225865; PubMed=6113690;  
 RA O'Shea M., Adams M.E.;  
 RT "Pentapeptide (proctolin) associated with an identified neuron.";  
 RL Science 213:567-569 (1981).  
 CC -!- FUNCTION: Stimulates cardiac output and hindgut motility,  
 CC modulates visceral and skeletal muscle in many arthropods.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Found in the lateral white neurons.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC -----  
 DR PIR; A01644; HOROHA.  
 SQ SEQUENCE 5 AA; 649 MW; 71B7673B4600000 CRC64;  
 KW Direct protein sequencing; Neuropeptide.  
 Query Match 55.6%; Score 5; DB 1; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 R 1  
 Db 1 R 1  
 RESULT 17  
 UF01 MOUSE  
 ID UF01 MOUSE STANDARD; PRT; 5 AA.  
 AC P38639;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Unknown protein from 2D-PAGE of fibroblasts (P19) (Fragment).  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 CC Muridae; Murinae; Mus.  
 CC NCBI\_TaxID=10090;  
 RN [1]  
 RP PROTEIN SEQUENCE.  
 RC TISSUE=Fibroblast;  
 RX MEDLINE=95009907; PubMed=7523108;  
 RA Merrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;  
 RT "Separation and sequencing of familial and novel murine proteins using  
 RT preparative two-dimensional gel electrophoresis.";  
 RL Electrophoresis 15:735-745 (1994).  
 CC -!- MISCELLANEOUS: On the 2D-gel the determined pI of this unknown  
 CC protein is: 6.6, its MW is: 19 kDa.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC -----  
 KW Direct protein sequencing.  
 FT NON\_TER 5  
 SQ SEQUENCE 5 AA; 717 MW; 7364087043100000 CRC64;  
 Query Match 55.6%; Score 5; DB 1; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 R 1  
 Db 4 R 4

RESULT 18  
 ACPH RABBIT  
 ID ACPH RABBIT STANDARD; PRT; 6 AA.  
 AC P25154;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 01-FEB-2005 (Rel. 46, Last annotation update)  
 DE Acylamino-acid-releasing enzyme (EC 3.4.19.1) (AARE) (Acyl-peptide  
 DE hydrolase) (APH) (Acylaminoacyl-peptidase) (Fragment).  
 GN Name=APRH;  
 OS Oryctolagus cuniculus (Rabbit).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Euarchontoglires; Glires; Lagomorpha; Leporidae;  
 CC Oryctolagus.  
 CC NCBI\_TaxID=9986;  
 RN [1]  
 RP PROTEIN SEQUENCE.  
 RC TISSUE=Muscle;  
 RX MEDLINE=92222120; PubMed=1807161;  
 RA Krishna R.G., Chin C.C.Q., Wold F.;  
 RT "N-terminal sequence analysis of N alpha-acetylated proteins after  
 RT unblocking with N-acylaminoacyl-peptide hydrolase.";  
 RL Anal. Biochem. 199:45-50 (1991).  
 CC -!- FUNCTION: This enzyme catalyzes the hydrolysis of the N-terminal  
 CC peptide bond of an N-acetylated peptide to generate an N-  
 CC acetylated amino acid and a peptide with a free N-terminus. It  
 CC preferentially cleaves off AC-Ala, Ac-Met and Ac-Ser.  
 CC -!- CATALYTIC ACTIVITY: Cleavage of an N-acetyl or N-formyl amino acid  
 CC from the N-terminus of a polypeptide.  
 CC -!- SUBUNIT: Homotetramer.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- SIMILARITY: Belongs to the peptidase S9C family.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC -----  
 DR PIR; A49792; A49792.  
 DR InterPro; IPR002471; Pept\_S9\_AS.  
 DR PROSITE; PS00708; PRO\_ENDOPEP\_SER; PARTIAL.  
 KW Acetylation; Direct protein sequencing; Hydrolase.  
 FT MOD\_RES 1 1 N-acetylmethionine.  
 FT NON\_TER 6 6  
 SQ SEQUENCE 6 AA; 775 MW; 6732D6C40B16F000 CRC64;  
 Query Match 55.6%; Score 5; DB 1; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 R 1  
 Db 3 R 3  
 RESULT 19  
 FARP MONEX  
 ID FARP MONEX STANDARD; PRT; 6 AA.  
 AC P41966;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE FMRamide-like neuropeptide GNFRF-amide.  
 OS Moniezia expansa (Sheep tapeworm).  
 CC Eukaryota; Metazoa; Platyhelminthes; Cestoda; Eucestoda;  
 CC Cyclophyllidae; Anoplocephalidae; Moniezia.  
 CC NCBI\_TaxID=28841;  
 RN [1]  
 RP PROTEIN SEQUENCE.  
 RX MEDLINE=93312289; PubMed=8323531;

RA Maule A.G., Shaw C., Halton D.W., Thim L.;  
RT "GNFPRamide: a novel FMRamide-immunoreactive peptide isolated from  
RL the sheep tapeworm, Moniezia expansa.";  
RL Biochem. Biophys. Res. Commun. 193:1054-1060(1993).  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- SIMILARITY: Belongs to the FARP (FMRamide related peptide)  
CC family.  
CC -----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC -----  
CC PIR; A43129; A43129.  
DR Amidation; Direct protein sequencing; Neuropeptide.  
KW MOD\_RES 6 Phenylalanine amide.  
FT MOD\_RES 6 787 MW; 69D409C9C4481000 CRC64;  
SQ SEQUENCE 6 AA; 787 MW; 69D409C9C4481000 CRC64;  
Query Match 55.6%; Score 5; DB 1; Length 6;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 R 1  
DB 5 R 5  
RESULT 20  
ID PFY1\_PENMO STANDARD; PRT; 6 AA.  
AC P84005;  
DT 05-JUL-2004 (Rel. 44, Created)  
DT 05-JUL-2004 (Rel. 44, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Peptide tyrosine phenylalanine 1 (Pem-PYF1).  
OS Penaeus monodon (Penaeid shrimp).  
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
OC Eumalacostraca; Eucarida; Decapoda; Dendrobranchiata; Penaeoidea;  
OC Penaeidae; Penaeus.  
OX NCBI\_TaxID=6687;  
[1]  
RN PROTEIN SEQUENCE, TISSUE SPECIFICITY, AND MASS SPECTROMETRY.  
RP TISSUE=Eyestalk;  
RC PubMed=12431727; DOI=10.1016/S0196-9781(02)00176-6;  
RX Sithigornkul P., Pupuem J., Krungkarn C., Longyant S., Panchan N.,  
RA Chaisuthangkura P., Sithigornkul W., Petsom A.;  
RT "Four novel PYFs: members of NP1/PP peptide superfamily from the  
RT eyestalk of the giant tiger prawn Penaeus monodon.";  
RL Peptides 23:1895-1906(2002).  
CC -!- FUNCTION: May act as a neurotransmitter, neuromodulator or  
CC neurohormone.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Limited to neuronal cell bodies, neuronal  
CC processes and sinus gland.  
CC -!- MASS SPECTROMETRY: MW=801.5; METHOD=MALDI; RANGE=1-6; NOTE=Ref.1.  
CC -!- SIMILARITY: Belongs to the NPY family.  
CC -----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC -----  
CC InterPro; IPR001955; Pancreatic\_hormn.  
DR PROSITE; PS00265; PANCREATIC\_HORMONE\_1; PARTIAL.  
DR PROSITE; PS0276; PANCREATIC\_HORMONE\_2; PARTIAL.  
KW Amidation; Direct protein sequencing; Neuropeptide.  
FT MOD\_RES 6 Phenylalanine amide (Potential).  
SQ SEQUENCE 6 AA; 802 MW; 69D417740DC46000 CRC64;  
Query Match 55.6%; Score 5; DB 1; Length 6;  
Best Local Similarity 100.0%; Pred. No. 0;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 R 1  
DB 1 R 1  
RESULT 21  
ID P82181\_SPIOL PRELIMINARY; PRT; 6 AA.  
AC P82181;  
DT 01-JUN-2000 (TrEMBLrel. 14, Created)  
DT 01-JUN-2000 (TrEMBLrel. 14, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Chloroplast 50S ribosomal protein L10 beta (Fragment).  
OS Spinacia oleracea (Spinach).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Caryophyllales; Amaranthaceae; Spinacia.  
OX NCBI\_TaxID=3562;  
[1]  
RN PROTEIN SEQUENCE.  
RP STRAIN=cv. ALVARO; TISSUE=Leaf;  
RC MEDLINE=20435798; PubMed=10874046; DOI=10.1074/jbc.M005012200;  
RX Yamaguchi K., Subramanian A.R.;  
RT "The plastid ribosomal proteins. Identification of all the proteins in  
RT the 50 S subunit of an organelle ribosome (chloroplast).";  
RL J. Biol. Chem. 275:28466-28482(2000).  
CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.  
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.  
CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.  
CC -!- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 kDa.  
CC -!- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.  
DR GO; GO:0009507; C:chloroplast; IEA.  
DR GO; GO:0019843; F:rRNA binding; IEA.  
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.  
DR InterPro; IPR002363; Ribosomal L10eub.  
DR PROSITE; PS01109; RIBOSOMAL\_L10; PARTIAL.  
KW Chloroplast; Ribosomal protein; rRNA-binding.  
FT NON\_TER 6  
SQ SEQUENCE 6 AA; 675 MW; 6321B415B05DB000 CRC64;  
Query Match 55.6%; Score 5; DB 2; Length 6;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 R 1  
DB 4 R 4  
RESULT 22  
ID P82182\_SPIOL PRELIMINARY; PRT; 6 AA.  
AC P82182;  
DT 01-JUN-2000 (TrEMBLrel. 14, Created)  
DT 01-JUN-2000 (TrEMBLrel. 14, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Chloroplast 50S ribosomal protein L10 gamma (Fragment).  
OS Spinacia oleracea (Spinach).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Caryophyllales; Amaranthaceae; Spinacia.  
OX NCBI\_TaxID=3562;  
[1]  
RN PROTEIN SEQUENCE.  
RP STRAIN=cv. ALVARO; TISSUE=Leaf;  
RC MEDLINE=20435798; PubMed=10874046; DOI=10.1074/jbc.M005012200;  
RX Yamaguchi K., Subramanian A.R.;  
RT "The plastid ribosomal proteins. Identification of all the proteins in  
RT the 50 S subunit of an organelle ribosome (chloroplast).";  
RL J. Biol. Chem. 275:28466-28482(2000).  
CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.

```

CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
CC -!- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 kDa.
CC -!- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
DR GO: GO:0009507; C:chloroplast; IEA.
DR GO: GO:0019843; F:rRNA binding; IEA.
DR GO: GO:0003735; F:structural constituent of ribosome; IEA.
DR InterPro: IPR002363; Ribosomal L10eub.
DR PROSITE: PS01109; RIBOSOMAL L10; PARTIAL.
KW Chloroplast; Ribosomal protein; rRNA-binding.
FT NON_TER 6
SQ SEQUENCE 6 AA; 675 MW; 6321B415B05DB000 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1
DB 4 R 4

RESULT 23
P82541_SPIOL
ID P82541_SPIOL PRELIMINARY; PRT; 6 AA.
AC P82541;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Chloroplast 30S ribosomal protein S19 beta (fragment).
OS Spinacia oleracea (Spinach).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC Caryophyllales; Amaranthaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP PROTEIN SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.
RC STRAIN=cv. ALVARO; TISSUE=Leaf.
RX MEDLINE=20435797; PubMed=10874039; DOI=10.1074/jbc.M004350200;
RA Yamaguchi K., von Knoblauch K., Subramanian A.R.;
RT "The plastid ribosomal proteins. Identification of all the proteins in
the 30S subunit of an organelle ribosome (chloroplast).";
RL J. Biol. Chem. 275:28455-28465(2000).
CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 16S RIBOSOMAL RNA.
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
CC -!- MASS SPECTROMETRY: MW=10477.0; METHOD=ELECTROSPRAY.
CC -!- MASS SPECTROMETRY: MW=10495; METHOD=MALDI.
CC -!- MISCELLANEOUS: S19 ALPHA AND BETA FORMS DIFFER IN PI. S19 BETA
FORM IS THE MINOR BASIC FORM.
CC -!- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 12 kDa.
CC -!- SIMILARITY: BELONGS TO THE S19P FAMILY OF RIBOSOMAL PROTEINS.
DR GO: GO:0009507; C:chloroplast; IEA.
DR GO: GO:0019843; F:rRNA binding; IEA.
DR GO: GO:0003735; F:structural constituent of ribosome; IEA.
DR InterPro: IPR002222; Ribosomal S19.
DR PROSITE: PS00323; RIBOSOMAL S19; PARTIAL.
KW Chloroplast; Ribosomal protein; rRNA-binding.
FT NON_TER 6
SQ SEQUENCE 6 AA; 732 MW; 63333735A411C000 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1
DB 2 R 2

RESULT 24
CAP6_CANAL
ID CAP6_CANAL STANDARD; PRT; 7 AA.
AC P83784;
DT 01-FEB-2005 (Rel. 46, Created)
DT 01-FEB-2005 (Rel. 46, Last sequence update)
DT 01-FEB-2005 (Rel. 46, Last annotation update)
DE Cytoplasmic antigenic protein 6 (fragment).
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP PROTEIN SEQUENCE, SUBCELLULAR LOCATION, AND ANTIGENICITY.
RC STRAIN=SC5314; TISSUE=Protoplast;
RX PubMed=15378761; DOI=10.1002/pmic.200400903;
RA Pitarich A., Abian J., Carrascal M., Sanchez M., Nombela C., Gil C.;
RT "Proteomics-based identification of novel Candida albicans antigens
for diagnosis of systemic candidiasis in patients with underlying
hematological malignancies.";
RT Proteomics 4:3084-3106(2004).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- MISCELLANEOUS: Has antigenic properties. Elicits a specific immune
response in systemic candidiasis human patients undergoing
malignant hematological disorders.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC -----
CC Antigen; Direct protein sequencing.
KW NON_TER 1
FT NON_TER 7
SQ SEQUENCE 7 AA; 900 MW; 740736C6D046DAC0 CRC64;

Query Match 55.6%; Score 5; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1
DB 7 R 7

RESULT 25
CARP_MYTED
ID CARP_MYTED STANDARD; PRT; 7 AA.
AC P10420;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Catch-relaxing peptide (CARP).
OS Mytilus edulis (Blue mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
OC Mytiloidea; Mytilidae; Mytilinae; Mytilus.
OX NCBI_TaxID=6550;
RN [1]
RP PROTEIN SEQUENCE.
RX MEDLINE=88052022; PubMed=3676797;
RA Hirata T., Kubota I., Takabatake I., Kawahara A., Shimamoto N.,
RA Muneoka Y.;
RT "Catch-relaxing peptide isolated from Mytilus pedal ganglia.";
RL Brain Res. 422:374-376(1987).
CC -!- FUNCTION: This peptide exhibits both potentiating (contraction)
and inhibitory (relaxation) effects on the anterior bysuss
retractor muscle.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC -----

```

```

DR PIR; A29342; ECMUCR.
KW Amidation; Direct protein sequencing; Hormone.
FT MOD_RES 7 Leucine amide.
SQ SEQUENCE 7 AA; 831 MW; 6734072687669DB0 CRC64;

Query Match 55.6%; Score 5; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1
Db 6 R 6

RESULT 26
CHOX_ALCSP STANDARD; PRT; 7 AA.
AC P16101;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Choline oxidase (BC 1.1.3.17) (Fragment).
OS Alcaligenes sp.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Alcaligenes.
OX NCBI_TaxID=512;
RN [1]
RP PROTEIN SEQUENCE.
RX MEDLINE=8106769; PubMed=6997283;
RA Ohta-Fukuyama M., Miyake Y., Emi S., Yamano T.;
RT "Identification and properties of the prosthetic group of choline
RT oxidase from Alcaligenes sp.";
RL J. Biochem. 88:197-203(1980).
CC -1- CATALYTIC ACTIVITY: Choline + O(2) = betaine aldehyde + H(2)O(2).
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
DR PIR; A15398; A15398.
KW Direct protein sequencing; Oxidoreductase.
FT NON_TER 7
SQ SEQUENCE 7 AA; 839 MW; 7415B1E457644AC0 CRC64;

Query Match 55.6%; Score 5; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1
Db 7 R 7

RESULT 27
FAF1_ASCSU STANDARD; PRT; 7 AA.
AC P31889;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE FMRFamide-like neuropeptide AF1.
OS Ascaris suum (pig roundworm) (Ascaris lumbricoidea).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
OC Ascarididae; Ascaris.
OX NCBI_TaxID=6253;
RN [1]
RP PROTEIN SEQUENCE.
RX MEDLINE=90180465; PubMed=2627377;
RA Cowden C., Stretton A.O.W., Davis R.E.;
RT "AF1, a sequenced bioactive neuropeptide isolated from the nematode
RT Ascaris suum.";

RL Neuron 2:1465-1473(1989).
CC -1- FUNCTION: Potent modulator of inhibitory motoneurons. Reduces the
CC input resistance and blocks slow oscillatory potentials in these
CC cells.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Found in the nerve cords and a variety of
CC ganglia particularly in the anterior regions.
CC -1- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)
CC family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
KW Amidation; Direct protein sequencing; Neuropeptide.
FT MOD_RES 7 Phenylalanine amide.
SQ SEQUENCE 7 AA; 953 MW; 69D40059CB144350 CRC64;

Query Match 55.6%; Score 5; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1
Db 6 R 6

RESULT 28
FAF2_ASCSU STANDARD; PRT; 7 AA.
AC P67879; P31890;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE FMRFamide-like neuropeptide AF2.
OS Ascaris suum (pig roundworm) (Ascaris lumbricoidea).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
OC Ascarididae; Ascaris.
OX NCBI_TaxID=6253;
RN [1]
RP PROTEIN SEQUENCE.
RX MEDLINE=93324431; PubMed=8332542; DOI=10.1016/0196-9781(93)90127-3;
RA Cowden C., Stretton A.O.W.;
RT "AF2, an Ascaris neuropeptide: isolation, sequence, and bioactivity.";
RL Peptides 14:423-430(1993).
CC -1- FUNCTION: Has effects on muscle tension.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Found in the nerve cords and a variety of
CC ganglia particularly in the anterior regions.
CC -1- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)
CC family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
KW Amidation; Direct protein sequencing; Neuropeptide.
FT MOD_RES 7 Phenylalanine amide.
SQ SEQUENCE 7 AA; 992 MW; 69D4073B5B1EE350 CRC64;

Query Match 55.6%; Score 5; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1
Db 6 R 6

```

## RESULT 29

FAP2\_PANRE STANDARD; PRT; 7 AA.  
 AC P67880; P31890;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE FMRFamide-like neuropeptide AP2.  
 OS Panagrellus redivivus.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;  
 OC Panagrolaimoidea; Panagrolaimidae; Panagrellus.  
 OX NCBI\_TaxID=6233;  
 RN [1]  
 RP PROTEIN SEQUENCE.  
 RX MEDLINE=95060998; PubMed=7970891;  
 RA Maule A.G., Shaw C., Bowman J.W.;  
 RT "The FMRFamide-like neuropeptide AP2 (*Ascaris suum*) is present in the  
 RT free-living nematode, *Panagrellus redivivus* (Nematoda, Rhabditida).";  
 RL Parasitology 109:351-356(1994).  
 CC -!- FUNCTION: Has effects on muscle tension.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Found in the nerve cords and a variety of  
 CC ganglia particularly in the anterior regions.  
 CC -!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)  
 CC family.  
 CC  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.

KW Amidation; Direct protein sequencing; Neuropeptide.  
 FT MOD\_RES 7 Phenylalanine amide.  
 SQ SEQUENCE 7 AA; 992 MW; 69D4073B5B1E350 CRC64;

Query Match 55.6%; Score 5; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 R 1  
 Db 6 R 6

## RESULT 30

FAR1\_HELTI STANDARD; PRT; 7 AA.  
 AC P41871;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE FMRFamide-like neuropeptide GDPFLRP-amide.  
 OS Helisoma trivolvis (Snail).  
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;  
 OC Lymnaeoidea; Planorbidae; Helisoma.  
 OX NCBI\_TaxID=27815;  
 RN [1]  
 RP PROTEIN SEQUENCE.  
 RX MEDLINE=94286417; PubMed=7912428; DOI=10.1016/0196-9781(94)90166-X;  
 RA Madrid K.P., Price D.A., Greenberg M.J., Khan H.R., Saleuddin A.S.M.;  
 RT "FMRFamide-related peptides from the kidney of the snail, *Helisoma*  
 RT trivolvis".  
 RL Peptides 15:31-36(1994).  
 CC -!- FUNCTION: Appears to be involved in osmoregulation by affecting  
 CC the kidney, mantle and skin.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Kidney, skin, mantle and the hemolymph.  
 CC -!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)  
 CC family.  
 CC  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.

KW Amidation; Direct protein sequencing; Neuropeptide.  
 FT MOD\_RES 7 Phenylalanine amide.  
 SQ SEQUENCE 7 AA; 851 MW; 69D40729D76AA810 CRC64;

Query Match 55.6%; Score 5; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 R 1  
 Db 6 R 6

## RESULT 31

FAR1\_MACRS STANDARD; PRT; 7 AA.  
 AC P83274;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE FMRFamide-like neuropeptide FLPI (DRNFLRF-amide).  
 OS Macrobrachium rosenbergii (Giant fresh water prawn).  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Caridea;  
 OC Palaemonoidea; Palaemonidae; Macrobrachium.  
 OX NCBI\_TaxID=79674;  
 RN [1]  
 RP PROTEIN SEQUENCE, AND MASS SPECTROMETRY.  
 RC TISSUE=Byestalk;  
 RA Sithigornkul P.; Saraitongkum W., Jaidechoey S., Longyant S.,  
 RA Sithigornkul W.;  
 RT "Novel FMRFamide-like neuropeptides from the eyestalk of the giant  
 RT freshwater prawn *Macrobrachium rosenbergii*.";  
 RL Comp. Biochem. Physiol. 120B:587-595(1998).  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- MASS SPECTROMETRY: MW=965.7; METHOD=MALDI; RANGE=1-7; NOTE=Ref.1.  
 CC -!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)  
 CC family.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.

DR GO; GO:0007218; P:neuropeptide signaling pathway; TAS.  
 KW Amidation; Direct protein sequencing; Neuropeptide.

FT MOD\_RES 7 Phenylalanine amide.  
 SQ SEQUENCE 7 AA; 967 MW; 69D40729C4540AC0 CRC64;

Query Match 55.6%; Score 5; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 R 1  
 Db 2 R 2

## RESULT 32

FAR1\_PROCL STANDARD; PRT; 7 AA.  
 AC P38499;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Cardioexcitatory FMRFamide homolog NF1.  
 OS *Procamburus clarkii* (Red swamp crayfish).

OC	Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;	FT	MOD_RES	7	7	Phenylalanine amide.
OC	Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidea;	SQ	SEQUENCE	7 AA; 967 MW; 69D40729C4540AC0 CRC64;		
OC	Astacidea; Cambaridae; Procambarus.					
OX	NCBI_TaxID=6728;					
RN	[1]					
RP	PROTEIN SEQUENCE.					
RC	TISSUE=Pericardial organs;					
RX	MEDLINE=93248032; PubMed=8387183; DOI=10.1016/0196-9781(93)90021-8;					
RA	Mercier A.J., Orchard I., Tebrugge V., Skerrett M.;					
RT	"Isolation of two FMRFamide-related peptides from crayfish pericardial organs.";					
RL	Peptides 14:137-143(1993).					
CC	-!- FUNCTION: Increases the rate and amplitude of spontaneous contractions of semi-isolated hearts. Increases the amplitude of excitatory postsynaptic potentials in abdominal extensor muscle.					
CC	-!- SUBCELLULAR LOCATION: Secreted.					
CC	-!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide) family.					
CC	-----					
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.					
CC	-----					
CC	Amidation; Direct protein sequencing; Neuropeptide.					
KW	MOD_RES 7 Phenylalanine amide.					
FT	SEQUENCE 7 AA; 966 MW; 69D40729C4540420 CRC64;					
SQ	SEQUENCE 7 AA; 966 MW; 69D40729C4540420 CRC64;					
	Query Match 55.6%; Score 5; DB 1; Length 7;					
	Best Local Similarity 100.0%; Pred. No. 0;					
	Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1 R 1					
Db	2 R 2					
	-----					
	RESULT 33					
FAR2	PROCL					
ID	FAR2_PROCL	STANDARD;	PRT;	7 AA.		
AC	F38498;					
DT	01-OCT-1994 (Rel. 30, Created)					
DT	01-OCT-1994 (Rel. 30, Last sequence update)					
DT	05-JUL-2004 (Rel. 44, Last annotation update)					
DE	Cardioexcitatory FMRFamide homolog DF2.					
OS	Procambarus clarkii (Red swamp crayfish).					
OC	Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;					
OC	Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidea;					
OC	Astacidea; Cambaridae; Procambarus.					
OX	NCBI_TaxID=6728;					
RN	[1]					
RP	PROTEIN SEQUENCE.					
RC	TISSUE=Pericardial organs;					
RX	MEDLINE=93248032; PubMed=8387183; DOI=10.1016/0196-9781(93)90021-8;					
RA	Mercier A.J., Orchard I., Tebrugge V., Skerrett M.;					
RT	"Isolation of two FMRFamide-related peptides from crayfish pericardial organs.";					
RL	Peptides 14:137-143(1993).					
CC	-!- FUNCTION: Increases the rate and amplitude of spontaneous contractions of semi-isolated hearts. Increases the amplitude of excitatory postsynaptic potentials in abdominal extensor muscle.					
CC	-!- SUBCELLULAR LOCATION: Secreted.					
CC	-!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide) family.					
CC	-----					
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.					
CC	-----					
CC	Amidation; Direct protein sequencing; Neuropeptide.					
KW	MOD_RES 7 Phenylalanine amide.					
FT	SEQUENCE 7 AA; 966 MW; 69D40729C4540420 CRC64;					
SQ	SEQUENCE 7 AA; 966 MW; 69D40729C4540420 CRC64;					
	Query Match 55.6%; Score 5; DB 1; Length 7;					
	Best Local Similarity 100.0%; Pred. No. 0;					
	Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1 R 1					
Db	2 R 2					
	-----					
	RESULT 34					
FAR3	HAECCO					
ID	FAR3_HAECCO	STANDARD;	PRT;	7 AA.		
AC	P81298;					
DT	30-MAY-2000 (Rel. 39, Created)					
DT	30-MAY-2000 (Rel. 39, Last sequence update)					
DT	05-JUL-2004 (Rel. 44, Last annotation update)					
DE	FMRFamide-like neuropeptide PF3 (KSAYMRF-amide).					
OS	Haemochus contortus (Barber pole worm).					
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;					
OC	Trichostrongyloidea; Haemonchidae; Haemonchinae; Haemonchus.					
OX	NCBI_TaxID=6289;					
RN	[1]					
RP	PROTEIN SEQUENCE.					
RC	TISSUE=Neuron;					
RX	MEDLINE=99318264; PubMed=10391380; DOI=10.1016/S0166-6851(99)00057-2;					
RA	Marks N.J., Sangster N.C., Maule A.G., Halton D.W., Thompson D.P.,					
RA	Geary T.G., Shaw C.;					
RT	"Structural characterisation and pharmacology of KHEYLRFamide (AF2) and KSAYMRFamide (PF3/AF8) from Haemonchus contortus.";					
RL	Mol. Biochem. Parasitol. 100:185-194(1999).					
CC	-!- FUNCTION: Active on neuromusculature.					
CC	-!- SUBCELLULAR LOCATION: Secreted.					
CC	-!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide) family.					
CC	-----					
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.					
CC	-----					
CC	Amidation; Direct protein sequencing; Neuropeptide.					
KW	MOD_RES 7 Phenylalanine amide.					
FT	SEQUENCE 7 AA; 902 MW; 69D4068B5DC5B350 CRC64;					
SQ	SEQUENCE 7 AA; 902 MW; 69D4068B5DC5B350 CRC64;					
	Query Match 55.6%; Score 5; DB 1; Length 7;					
	Best Local Similarity 100.0%; Pred. No. 0;					
	Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1 R 1					
Db	6 R 6					
	-----					
	RESULT 35					
FAR3	PANRE					
ID	FAR3_PANRE	STANDARD;	PRT;	7 AA.		
AC	P41874;					
DT	01-NOV-1995 (Rel. 32, Created)					
DT	01-NOV-1995 (Rel. 32, Last sequence update)					
DT	05-JUL-2004 (Rel. 44, Last annotation update)					
DE	FMRFamide-like neuropeptide PF3 (KSAYMRF-amide).					
OS	Panagrellus redivivus.					
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;					
OC	Panagrolaimoidea; Panagrolaimidae; Panagrellus.					
OX	NCBI_TaxID=6233;					
RN	[1]					
RP	PROTEIN SEQUENCE. AND SYNTHESIS.					
RX	MEDLINE=94235053; PubMed=8179635;					
RA	Maule A.G., Shaw C., Bowman J.W., Halton D.W., Thompson D.P.,					

RA Geary T.G., Thim L.;  
RT "KSAFMRamide: a novel FMRamide-related heptapeptide from the free-living nematode, *Panagrellus redivivus*, which is myoactive in the parasitic nematode, *Ascaris suum*.";  
RL Biochem. Biophys. Res. Commun. 200:973-980(1994).  
CC !- FUNCTION: Myoactive; induces a rapid concentration-dependent muscle tension increase.  
CC !- SUBCELLULAR LOCATION: Secreted.  
CC !- SIMILARITY: Belongs to the FARP (FMRamide related peptide) family.  
CC  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.  
CC  
CC PIR; PC2132; PC2132.  
KW Amidation; Direct protein sequencing; Neuropeptide.  
FT MOD\_RES 7 Phenylalanine amide.  
SQ SEQUENCE 7 AA; 902 MW; 69D406B5DC5B350 CRC64;  
Query Match 55.6%; Score 5; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 R 1  
Db 6 R 6  
RESULT 36  
FAR4 PANRE  
ID FAR4 PANRE STANDARD; PRT; 7 AA.  
AC P41875;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DE FMRamide-like neuropeptide PF4 (KPNFIRF-amide).  
OS *Panagrellus redivivus*.  
CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;  
CC Panagrolaimoidea; Panagrolaimidae; Panagrellus.  
OX NCBI\_TaxID=6233;  
RN [1]  
PP  
RP PROTEIN SEQUENCE, AND SYNTHESIS.  
RX MEDLINE=95232026; PubMed=7716079; DOI=10.1016/0196-9781(94)00162-Y;  
RA Maule A.G., Shaw C., Bowman J.W., Halton D.W., Thompson D.P., Thim L., Kubiak T.M., Martin R.A., Geary T.G.;  
RT "Isolation and preliminary biological characterization of KPNFIRFamide, a novel FMRamide-related peptide from the free-living nematode, *Panagrellus redivivus*.";  
RL Peptides 16:87-93(1995).  
CC !- FUNCTION: Myoactive; induces a rapid concentration-dependent muscle tension increase.  
CC !- SUBCELLULAR LOCATION: Secreted.  
CC !- SIMILARITY: Belongs to the FARP (FMRamide related peptide) family.  
CC  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.  
CC  
CC Amidation; Direct protein sequencing; Neuropeptide.  
KW MOD\_RES 7 Phenylalanine amide.  
SQ SEQUENCE 7 AA; 921 MW; 69D40059C4576350 CRC64;  
Query Match 55.6%; Score 5; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 R 1

Db 6 R 6  
RESULT 37  
PAR5 HIRME  
ID PAR5 HIRME STANDARD; PRT; 7 AA.  
AC P42564;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DE FMRamide-like neuropeptide GGYMRP-amide.  
OS *Hirudo medicinalis* (Medicinal leech).  
CC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;  
CC Arhynchobdellida; Hirudiniformes; Hirudinidae; Hirudo.  
OX NCBI\_TaxID=6421;  
RN [1]  
PP  
RP PROTEIN SEQUENCE.  
RX MEDLINE=92195954; PubMed=1686933; DOI=10.1016/0196-9781(91)90035-N;  
RA Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;  
RT "Identification of RFamide neuropeptides in the medicinal leech.";  
RL Peptides 12:897-908(1991).  
CC !- SUBCELLULAR LOCATION: Secreted.  
CC !- SIMILARITY: Belongs to the FARP (FMRamide related peptide) family.  
CC  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.  
CC  
CC Amidation; Direct protein sequencing; Neuropeptide.  
FT MOD\_RES 7 Phenylalanine amide.  
SQ SEQUENCE 7 AA; 858 MW; 69D406B853387810 CRC64;  
Query Match 55.6%; Score 5; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 R 1  
Db 6 R 6  
RESULT 38  
FARB CALVO  
ID FARB CALVO STANDARD; PRT; 7 AA.  
AC P41866;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DE CalliFMRamide 11.  
OS *Calliphora vomitoria* (Blue blowfly).  
CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;  
CC Calliphoridae; Calliphora.  
OX NCBI\_TaxID=27454;  
RN [1]  
PP  
RP PROTEIN SEQUENCE.  
RX TISSUE=Thoracic ganglion;  
RX MEDLINE=92196111; PubMed=1549595;  
RA Duve H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I., Rehfeld J.F., Thorpe A.;  
RT "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2 neuropeptides (designated calliFMRamides) from the blowfly *Calliphora vomitoria*.";  
RL Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).  
CC !- SUBCELLULAR LOCATION: Secreted.  
CC !- SIMILARITY: Belongs to the FARP (FMRamide related peptide) family.  
CC  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC -----

DR PIR; B44787; B44787.  
KW Amidation; Direct protein sequencing; Neuropeptide.  
FT MOD RES 7 7 Phenylalanine amide.  
SQ SEQUENCE 7 AA; 926 MW; 69D40699C44AB700 CRC64;

Query Match 55.6%; Score 5; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1  
|  
Db 6 R 6

RESULT 39  
ID HCVB CONCC STANDARD; PRT; 7 AA.  
AC P84620;  
DT 13-SEP-2005 (Rel. 48, Created)  
DT 13-SEP-2005 (Rel. 48, Last sequence update)  
DT 13-SEP-2005 (Rel. 48, Last annotation update)  
DE Hemocyanin subunit B (CCH-B) (Fragment).  
OS Concholepas concholepas (Barnacle rock-shell).  
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;  
OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;  
OC Neogastropoda; Muricoidae; Muricidae; Concholepas.  
OX NCBI\_TaxID=137544;  
[1]  
RN PROTEIN SEQUENCE, COFACTOR, SUBUNIT, SUBCELLULAR LOCATION, AND TISSUE  
RP SPECIFICITY.  
RC TISSUE=Hemolymph;  
RX PubMed=15075320; DOI=10.1074/jbc.M400903200;  
RA De Ioannes P., Molledo B., Oliva H., Facheo R., Faunes F.,  
RA De Ioannes A.E., Becker M.I.;  
RT "Hemocyanin of the molluscan Concholepas concholepas exhibits an  
RT unusual heterodecameric array of subunits.";  
RL J. Biol. Chem. 279:26134-26142(2004).  
CC -!- FUNCTION: Hemocyanins are copper-containing oxygen carriers  
CC occurring freely dissolved in the hemolymph of many mollusks and  
CC arthropods.  
CC -!- COFACTOR: Binds 2 copper ions per functional unit.  
CC -!- SUBUNIT: Heterododecamer composed of A and B subunits, each  
CC containing 8 globular oxygen-binding functional units.  
CC Heterogenous decameric or multidecameric structures may also be  
CC formed.  
CC -!- SUBCELLULAR LOCATION: Extracellular.  
CC -!- TISSUE SPECIFICITY: Hemolymph.  
CC -!- PM: Forms a thioether bond between 2 amino acids (By similarity).  
CC -!- SIMILARITY: Belongs to the tyrosinase family. Hemocyanin  
CC subfamily.

-----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC -----  
DR PROSITE; PS00209; HEMOCYANIN\_1; PARTIAL.  
DR PROSITE; PS00210; HEMOCYANIN\_2; PARTIAL.  
DR PROSITE; PS00497; TYROSINASE\_1; PARTIAL.  
DR PROSITE; PS00498; TYROSINASE\_2; PARTIAL.  
KW Copper; Direct protein sequencing; Hemolymph; Metal-binding;  
KW Oxygen transport; Thioether bond; Transport.  
FT NON TER 7  
SQ SEQUENCE 7 AA; 855 MW; 6AB2C443241AE740 CRC64;  
Query Match 55.6%; Score 5; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 0;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 R 1  
|  
Db 3 R 3

RESULT 40  
ID IPVR CANAL STANDARD; PRT; 7 AA.  
AC P83777;  
DT 01-FEB-2005 (Rel. 46, Created)  
DT 01-FEB-2005 (Rel. 46, Last sequence update)  
DT 01-FEB-2005 (Rel. 46, Last annotation update)  
DE Inorganic pyrophosphatase (EC 3.6.1.1) (Pyrophosphate phospho-  
DE hydrolase) (PPase) (Fragment).  
GN Name=IPPI;  
OS Candida albicans (Yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.  
OX NCBI\_TaxID=5476;  
[1]

RN PROTEIN SEQUENCE, SUBCELLULAR LOCATION, AND ANTIGENICITY.  
RP STRAIN=SC5314; TISSUE=Protoplast; 200400903;  
RX PubMed=15378761; DOI=10.1002/pmic.200400903;  
RA Pitarch A., Abian J., Carrascal M., Sanchez M., Nombela C., Gil C.;  
RT "Proteomics-based identification of novel Candida albicans antigens  
RT for diagnosis of systemic candidiasis in patients with underlying  
RT hematological malignancies";  
RL Proteomics 4:3084-3106(2004).  
CC -!- CATALYTIC ACTIVITY: Diphosphate + H(2)O = 2 phosphate.  
CC -!- COFACTOR: Binds 4 magnesium ions per subunit. Other metal ions can  
CC support activity, but at a lower rate. Two magnesium ions are  
CC required for the activation of the enzyme and are present before  
CC substrate binds, two additional magnesium ions form complexes with  
CC substrate and product (By similarity).  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -!- MISCELLANEOUS: Has antigenic properties. Elicits a specific immune  
CC response in systemic candidiasis human patients undergoing  
CC malignant hematological disorders.  
CC -!- SIMILARITY: Belongs to the PPase family.

-----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC -----  
DR InterPro; IPR008162; Pyrophosphatase.  
DR PROSITE; PS00387; PPASE; PARTIAL.  
KW Antigen; Direct protein sequencing; Hydrolase; Magnesium;  
KW Metal-binding.  
FT NON TER 1 1  
FT NON TER 7 7  
SQ SEQUENCE 7 AA; 923 MW; 7409D37B1451ADB0 CRC64;

Query Match 55.6%; Score 5; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1  
|  
Db 7 R 7

RESULT 41  
ID LANC CARUI STANDARD; PRT; 7 AA.  
AC P36950;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 13-SEP-2005 (Rel. 46, Last annotation update)  
DE Lantibiotic carnocin UI49 (Fragment).



OS Carnobacterium sp. (strain UI49).  
OC Bacteria; Firmicutes; Lactobacillales; Carnobacteriaceae;  
OC Carnobacterium.  
OX NCBI\_TaxID=35782;  
RN [1]  
RP PROTEIN SEQUENCE.  
RX MEDLINE=92321768; PubMed=16222206;  
RA Stoffels G., Nissen-Meyer J., Gudmundsdottir A., Sletten K., Holo H.,  
RA Nes I.F.;  
RT "Purification and characterization of a new bacteriocin isolated from  
RT a Carnobacterium sp.";  
RL Appl. Environ. Microbiol. 58:1417-1422(1992).  
CC -!- FUNCTION: Lanthionine-containing peptide antibiotic (lantibiotic).  
CC Active on Gram-positive bacteria.  
CC -----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC -----  
CC Antibiotic; Antimicrobial; Bacteriocin; Direct protein sequencing;  
KW Lantibiotic.  
KW Lantibiotic.  
FT NON\_TER 7 7  
SQ SEQUENCE 7 AA; 786 MW; 741776D05B05B810 CRC64;  
  
Query Match 55.6%; Score 5; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 R 1  
|  
Db 7 R 7  
  
RESULT 42  
TY51\_LITRU  
ID TY51\_LITRU STANDARD; PRT; 7 AA.  
AC P82065;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 01-FEB-2005 (Rel. 46, Last annotation update)  
DE Tryptophyllin-5.1.  
OS Litoria rubella (Desert tree frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Neobatrachia; Hylidae; Hylidae;  
OC Pelodyadinae; Litoria.  
OX NCBI\_TaxID=104895;  
RN [1]  
RP PROTEIN SEQUENCE, AND MASS SPECTROMETRY.  
RC TISSUE=Skin secretion;  
RA Steinborner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C.,  
RA Tyler M.J., Wallace J.C.;  
RT "The structure of new peptides from the Australian red tree frog  
RT 'Litoria rubella'. The skin peptide profile as a probe for the study  
RT of evolutionary trends of amphibians.";  
RL Aust. J. Chem. 49:955-963(1996).  
CC -!- FUNCTION: May act as a neuromodulator or neurotransmitter.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Expressed by the skin dorsal glands.  
CC -!- MASS SPECTROMETRY: MW=965; METHOD=FAB; RANGE=1-7; NOTE=Ref.1.  
CC -----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC -----  
CC Amidation; Amphibian defense peptide; Direct protein sequencing;  
KW Neuropeptide; Pyrrolidone carboxylic acid.  
FT MOD\_RES 1 1 Pyrrolidone carboxylic acid.  
FT MOD\_RES 7 7 Arginine amide.  
SQ SEQUENCE 7 AA; 983 MW; 7401E9D3676046B0 CRC64;

Query Match 55.6%; Score 5; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 R 1  
|  
Db 7 R 7  
  
RESULT 43  
UF04\_MOUSE  
ID UF04\_MOUSE STANDARD; PRT; 7 AA.  
AC P38642;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Unknown protein from 2D-PAGE of fibroblasts (P46) (Fragment).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muridea; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP PROTEIN SEQUENCE.  
RC TISSUE=Fibroblast;  
RX MEDLINE=95009907; PubMed=7523108;  
RA Merrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;  
RT "Separation and sequencing of familial and novel murine proteins using  
RT preparative two-dimensional gel electrophoresis.";  
RL Electrophoresis 15:735-745(1994).  
CC -!- MISCELLANEOUS: On the 2D-gel the determined pI of this unknown  
CC protein is: 5.0, its MW is: 46 kDa.  
CC -----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC -----  
CC Direct protein sequencing.  
FT NON\_TER 7 7  
SQ SEQUENCE 7 AA; 766 MW; 68640AB77632700 CRC64;  
  
Query Match 55.6%; Score 5; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 R 1  
|  
Db 6 R 6  
  
RESULT 44  
UH11\_RAT  
ID UH11\_RAT STANDARD; PRT; 7 AA.  
AC P56576;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Unknown protein from 2D-PAGE of heart tissue (Spot P11) (Fragment).  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muridea; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP PROTEIN SEQUENCE.  
RC STRAIN=Wistar; TISSUE=Heart;  
RA Li X.-P., Pleissner K.-P., Scheler C., Regitz-Zagrosek V., Salikov J.,  
RA Jungblut P.R.;  
RL Submitted (SEP-1998) to Swiss-Prot.  
CC -!- MISCELLANEOUS: On the 2D-gel the determined pI of this unknown

```
CC      protein is: 8.5, its MW is: 42 kDa.
CC      -----
CC      This Swiss-Prot entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use as long as its content is in no way modified and this statement is not
CC      removed.
CC      -----
CC      Direct protein sequencing.
CC      UNSURE      2      7      S or A.
FT      NON_TER      2      7
SQ      SEQUENCE      7 AA; 775 MW; 6866DB040DC5A6B0 CRC64;

      Query Match      55.6%; Score 5; DB 1; Length 7;
      Best Local Similarity 100.0%; Pred. No. 0;
      Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 R 1
Db      4 R 4

RESULT 45
WWAL_ACHFUFU      STANDARD;      PRT;      7 AA.
AC      P35919;
DT      01-JUN-1994 (Rel. 29, Created)
DT      01-JUN-1994 (Rel. 29, Last sequence update)
DT      05-JUL-2004 (Rel. 44, Last annotation update)
DE      Wamide-1.
OS      Achatina fulica (Giant African snail).
OC      Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
OC      Sigurethra; Achatinoidea; Achatinidae; Achatina.
OX      NCBI_TaxID=6530;
RN      [1]
RP      PROTEIN SEQUENCE.
RC      TISSUE=Ganglion;
RX      MEDLINE=93265912; PubMed=8495720; DOI=10.1016/0014-5793(93)81458-C;
RA      Minakata H., Ikeda T., Muneoka Y., Kobayashi M., Nomoto K.;
RT      "Wamide-1, -2 and -3: novel neuromodulatory peptides isolated from
RT      ganglia of the African giant snail, Achatina fulica."
RL      FEBS Lett. 323:104-108(1993).
CC      -!- FUNCTION: Exhibits modulatory effects on the peripheral nervous
CC      system. Inhibits activity on a central neuron.
CC      -----
CC      This Swiss-Prot entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use as long as its content is in no way modified and this statement is not
CC      removed.
CC      -----
CC      PIR; S33245; S33245.
KW      Amidation; Direct protein sequencing; Neuropeptide.
FT      MOD_RES      7      7      Tryptophan amide.
SQ      SEQUENCE      7 AA; 993 MW; 7362D5B69B041310 CRC64;

      Query Match      55.6%; Score 5; DB 1; Length 7;
      Best Local Similarity 100.0%; Pred. No. 0;
      Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 R 1
Db      2 R 2

RESULT 46
Q95945 YEAST
ID      Q95945 YEAST PRELIMINARY;      PRT;      7 AA.
AC      Q95945;
DT      01-FEB-1997 (TrEMBLrel. 02, Created)
DT      01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT      01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE      Inside intron 5 (Fragment).
```

```
OS      Saccharomyces cerevisiae (Baker's yeast).
OG      Mitochondrion.
OC      Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC      Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX      NCBI_TaxID=4932;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=D273-10B;
RX      MEDLINE=81069885; PubMed=6254986;
RA      Bonitz S.G., Coruzzi G., Thalenfeld B.E., Tzagoloff A., Macino G.;
RT      "Assembly of the mitochondrial membrane system. Structure and
RT      nucleotide sequence of the gene coding for subunit 1 of yeast
RT      cytochrome oxidase."
RL      J. Biol. Chem. 255:11927-11941(1980).
DR      EMBL; V00694; CAA24066.1; -; Genomic.DNA.
DR      GO; GO:0005739; C:mitochondrion; IEA.
KW      Mitochondrion.
FT      NON_TER      1
SQ      SEQUENCE      7 AA; 859 MW; 75B7232362CDC460 CRC64;

      Query Match      55.6%; Score 5; DB 2; Length 7;
      Best Local Similarity 100.0%; Pred. No. 0;
      Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 R 1
Db      1 R 1

RESULT 47
Q15903 HUMAN
ID      Q15903 HUMAN PRELIMINARY;      PRT;      7 AA.
AC      Q15903;
DT      01-NOV-1996 (TrEMBLrel. 01, Created)
DT      01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT      01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE      Homo sapiens (clone XP7E7B) (Fragment).
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OX      Homo.
OC      NCBI_TaxID=9606;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      TISSUE=Placenta;
RA      Lee C.-C., Yazdani A., Wehnert M., Bailey J., Couch L., Xiong M.,
RA      Coolbaugh M.I., Chinault C.A., Baldini A., Lindsey E.A., Zhao Z.-Y.,
RA      Caskey C.T.H.;
RT      "Isolation of chromosome-specific genes by reciprocal probing of
RT      arrayed cDNAs and cosmid libraries."
RL      Hum. Mol. Genet. 0:0-0(1995).
DR      EMBL; L32082; AAA73893.1; -; mRNA.
FT      NON_TER      1
FT      NON_TER      7
SQ      SEQUENCE      7 AA; 849 MW; 6B040339CDD33DB0 CRC64;

      Query Match      55.6%; Score 5; DB 2; Length 7;
      Best Local Similarity 100.0%; Pred. No. 0;
      Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 R 1
Db      6 R 6

RESULT 48
Q8NHH7 HUMAN
ID      Q8NHH7 HUMAN PRELIMINARY;      PRT;      7 AA.
AC      Q8NHH7;
DT      01-OCT-2002 (TrEMBLrel. 22, Created)
DT      01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT      01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE      Mini-cistron.
```

```

GN Name=NHE3;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP NUCLEOTIDE SEQUENCE.
RA Malakooti J., Ramaswamy K.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF282824; AAM53436.1; -; Genomic DNA.
SQ SEQUENCE 7 AA; 842 MW; 74072DC772D406F0 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1
Db 2 R 2

RESULT 49
Q8TAQ4 HUMAN PRELIMINARY; PRT; 7 AA.
ID Q8TAQ4;
AC Q8TAQ4;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE YAP1 protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Uterus;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Klausner R.L., Feingold E.A., Grouse L.H., Deige J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buecow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Uterus;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC026212; AAH26212.2; -; mRNA.
SQ SEQUENCE 7 AA; 848 MW; 6AB2D1B6C2D406F0 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1
Db 2 R 2

```

```

Db 2 R 2

RESULT 50
O98866 SPIOL PRELIMINARY; PRT; 7 AA.
ID O98866;
AC O98866;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytochrome b/f subunit IV (Fragment).
OS Spinacia oleracea (Spinach).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Amaranthaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]_TaxID=3562;
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=86120353; PubMed=3003688;
RA Sijben-Mueller G., Hallick R.B., Alt J., Westhoff P., Herrmann R.G.;
RT "Spinach plastid genes coding for initiation factor IF-1, ribosomal
RT protein S11 and RNA polymerase alpha-subunit.";
RL Nucleic Acids Res. 14:1029-1044(1986).
DR EMBL; X03496; CAA27215.1; -; Genomic DNA.
DR GO; GO:0009507; C:chloroplast; IEA.
KW Chloroplast.
FT NON_TER
SQ SEQUENCE 7 AA; 907 MW; 644729D77409C420 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 R 1
Db 3 R 3

Search completed: January 25, 2006, 18:41:14
Job time : 92 secs

```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:32:42 ; Search time 78.5 Seconds

(without alignments)  
27.986 Million cell updates/sec

Title: US-10-771-242-295

Perfect score: 21

Sequence: 1 RRLNX 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :

- 1: Geneseqp21.\*
- 2: Geneseqp1980s.\*
- 3: Geneseqp1990s.\*
- 4: Geneseqp2000s.\*
- 5: Geneseqp2001s.\*
- 6: Geneseqp2002s.\*
- 7: Geneseqp2003as.\*
- 8: Geneseqp2003bs.\*
- 9: Geneseqp2004s.\*
- 9: Geneseqp2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	95.2	5	ADZ71770	Adz71770 p21-deriv
2	20	95.2	5	ADZ72025	Adz72025 p21-deriv
3	20	95.2	5	ADZ71763	Adz71763 p21-deriv
4	20	95.2	5	ADZ72027	Adz72027 p21-deriv
5	20	95.2	5	ADZ72107	Adz72107 p21-deriv
6	20	95.2	5	ADZ71617	Adz71617 p21-deriv
7	20	95.2	5	ADZ71663	Adz71663 p21-deriv
8	20	95.2	5	ADZ71764	Adz71764 p21-deriv
9	20	95.2	5	ADZ71765	Adz71765 p21-deriv
10	20	95.2	5	ADZ71746	Adz71746 p21-deriv
11	20	95.2	5	ADZ71618	Adz71618 p21-deriv
12	20	95.2	5	ADZ71662	Adz71662 p21-deriv
13	20	95.2	5	ADZ71664	Adz71664 p21-deriv
14	20	95.2	5	ADZ72026	Adz72026 p21-deriv
15	20	95.2	5	ADZ72106	Adz72106 p21-deriv
16	20	95.2	5	ADZ71747	Adz71747 p21-deriv
17	20	95.2	5	ADZ71748	Adz71748 p21-deriv
18	20	95.2	6	AAM56904	Aam56904 Enzyme in
19	20	95.2	7	AAM46084	Aam46084 H11 bindi
20	20	95.2	7	AAM46173	Aam46173 H11 bindi
21	20	95.2	7	AAM44565	Aam44565 H11 bindi
22	20	95.2	7	AAM44575	Aam44575 H11 bindi
23	20	95.2	7	AAM43976	Aam43976 H11 bindi
24	20	95.2	7	AAM44570	Aam44570 H11 bindi

25	20	95.2	7	4	AAM46480	Aam46480 H11 bindi
26	20	95.2	7	4	AAM44560	Aam44560 H11 bindi
27	20	95.2	7	4	AAM45974	Aam45974 H11 bindi
28	20	95.2	7	4	AAM46528	Aam46528 H11 bindi
29	20	95.2	7	4	AAM46485	Aam46485 H11 bindi
30	20	95.2	7	4	AAM45665	Aam45665 H11 bindi
31	20	95.2	7	4	AAM46523	Aam46523 H11 bindi
32	20	95.2	7	4	AAM46526	Aam46526 H11 bindi
33	20	95.2	8	2	AAW57007	Aaw57007 Enzyme in
34	20	95.2	8	4	AAU05744	Aau05744 p21 C-ter
35	20	95.2	8	9	ADZ71972	Adz71972 p21-deriv
36	20	95.2	8	9	ADZ71569	Adz71569 p21-deriv
37	20	95.2	9	5	AAE18742	Aae18742 Human leu
38	20	95.2	9	5	AAE18741	Aae18741 Human leu
39	20	95.2	9	5	AAE18780	Aae18780 Human leu
40	20	95.2	11	4	ABW01200	Abw01200 Saccharom
41	20	95.2	11	7	ABW01200	Abw01200 Saccharom
42	20	95.2	11	8	ADP44126	Adp44126 Yeast tra
43	20	95.2	12	4	AB599999	Ab599999 Internall
44	20	95.2	12	7	ADA88834	Ada88834 Internall
45	20	95.2	12	8	ADQ15921	Adq15921 Human Mas
46	20	95.2	14	2	AAR47006	Aar47006 LAR prote
47	20	95.2	15	2	AAW42267	Aaw42267 Biotinyla
48	20	95.2	15	2	AAW15509	Aaw15509 Peptide M
49	20	95.2	15	4	AB576940	Ab576940 Rat VG51-
50	20	95.2	15	5	AAE18774	Aae18774 Human leu
51	20	95.2	19	5	AAU93571	Aau93571 Granulocy
52	20	95.2	20	6	ADA08204	Ada08204 Human CRK
53	20	95.2	20	7	ADH47647	Adh47647 Human lun
54	20	95.2	20	7	ADH47648	Adh47648 Human lun
55	20	95.2	20	8	ADJ21567	Adj21567 Human lun
56	20	95.2	20	8	ADJ21566	Adj21566 Human lun
57	20	95.2	24	2	AAR85034	Aar85034 Peptide r
58	20	95.2	28	2	AAZ20367	Aaz20367 Human mic
59	20	95.2	32	2	AAR85789	Aar85789 Peptide r
60	20	95.2	33	2	AAZ48432	Aaz48432 Human pro
61	20	95.2	33	7	ABW79022	Abw79022 BK channe
62	20	95.2	34	3	AB34526	Ab34526 Human sec
63	20	95.2	35	4	ABB40086	Abb40086 Peptide #
64	20	95.2	35	4	AAW73524	Aaw73524 Human bon
65	20	95.2	35	4	AAW60842	Aaw60842 Human bra
66	20	95.2	35	4	ABG55251	Abg55251 Human liv
67	20	95.2	35	5	ABG54385	Abg43385 Human pep
68	20	95.2	36	2	AAZ59938	Aaz59938 Human myo
69	20	95.2	37	2	AAR85022	Aar85022 Peptide r
70	20	95.2	37	2	AAR85022	Aar85022 Peptide r
71	20	95.2	40	7	ABR82177	AbR82177 Source 27
72	20	95.2	40	7	ABR82163	AbR82163 Source 13
73	20	95.2	41	4	ABB03590	Abb03590 Human mus
74	20	95.2	41	6	ABU12884	Abu12884 Novel mus
75	20	95.2	41	8	ADJ28910	Adj28910 Human mus
76	20	95.2	47	3	AAZ45682	Aaz45682 Arabidops
77	20	95.2	47	3	ABE17914	Abb17914 Human ner
78	20	95.2	47	9	ADY66570	Ady66570 S. mansoni
79	20	95.2	48	8	ADZ07330	Adz07330 Staphyloc
80	20	95.2	49	4	AAZ23505	Aam23505 Human EST
81	20	95.2	49	9	ADY30821	Ady30821 Thale cre
82	20	95.2	50	4	ABZ42086	Abz42086 Peptide #
83	20	95.2	50	4	AAZ35888	Aam35888 Peptide #
84	20	95.2	50	4	AAW75778	Aaw75778 Human bon
85	20	95.2	50	4	AAU46056	Aau46056 Propionib
86	20	95.2	50	4	AAW62966	Aam62966 Human bra
87	20	95.2	50	4	ABG57516	Abg57516 Human liv
88	20	95.2	50	5	ABG45252	Abg45252 Human pep
89	20	95.2	50	6	ABM42575	Abm42575 Propionib
90	20	95.2	51	4	ABB32056	Abb32056 Peptide #
91	20	95.2	51	4	ABB22594	Abb22594 Protein #
92	20	95.2	51	4	AAU60742	Aau60742 Propionib
93	20	95.2	51	4	ABG52125	Abg52125 Human liv
94	20	95.2	51	5	ABP02714	Abp02714 Human ORF
95	20	95.2	51	5	ABG40076	Abg40076 Human pep
96	20	95.2	51	6	ABM57261	Abm57261 Propionib
97	20	95.2	52	8	ADS07331	AdS07331 Staphyloc

98	20	95.2	53	4	AAU47841	AAU47841	Human	imm	171	20	95.2	71	6	ABM51501	ABM51501	Propionib
99	20	95.2	53	4	AAU47841	AAU47841	Human	imm	172	20	95.2	72	5	ABP31506	ABP31506	Human ORF
100	20	95.2	53	6	ABM43360	ABM43360	Propionib		173	20	95.2	72	8	ADL04327	ADL04327	M. catarr
101	20	95.2	53	8	ADU66980	ADU66980	Human	sec	174	20	95.2	73	3	ABM14406	ABM14406	Human lip
102	20	95.2	54	3	AAU03645	AAU03645	Human	sec	175	20	95.2	73	4	ABM11006	ABM11006	Human rev
103	20	95.2	54	4	AAU46799	AAU46799	Propionib		176	20	95.2	73	4	ABM03867	ABM03867	Human gen
104	20	95.2	54	4	AAU47820	AAU47820	Propionib		177	20	95.2	73	5	ABG64571	ABG64571	Human alb
105	20	95.2	54	6	ABM44339	ABM44339	Propionib		178	20	95.2	73	5	ABG00887	ABG00887	Enterohae
106	20	95.2	54	6	ABM43318	ABM43318	Propionib		179	20	95.2	73	7	ADC00887	ADC00887	Enterohae
107	20	95.2	55	4	AAU060505	AAU060505	Propionib		180	20	95.2	73	7	ADC00210	ADC00210	Enterohae
108	20	95.2	55	6	ABM57024	ABM57024	Propionib		181	20	95.2	73	7	ADC00577	ADC00577	Enterohae
109	20	95.2	56	6	ABM20689	ABM20689	Peptide #		182	20	95.2	73	7	ADL77838	ADL77838	Albumin f
110	20	95.2	56	4	ABM42078	ABM42078	Peptide #		183	20	95.2	74	3	AAU74684	AAU74684	Neisseria
111	20	95.2	56	4	ABM42078	ABM42078	Peptide #		184	20	95.2	74	6	ABU70734	ABU70734	Human adi
112	20	95.2	56	4	ABM35880	ABM35880	Peptide #		185	20	95.2	75	5	ABP29218	ABP29218	Streptoco
113	20	95.2	56	4	ABM25675	ABM25675	Protein #		186	20	95.2	75	6	ABD10678	ABD10678	Allioco
114	20	95.2	56	4	ABM75770	ABM75770	Human bon		187	20	95.2	76	3	AAU14404	AAU14404	Rat Rhoga
115	20	95.2	56	4	AAU59928	AAU59928	Propionib		188	20	95.2	76	3	AAU38481	AAU38481	Peptide #
116	20	95.2	56	4	ABM62958	ABM62958	Human bra		189	20	95.2	77	4	AAU65585	AAU65585	Human bra
117	20	95.2	56	4	ABG57510	ABG57510	Human liv		190	20	95.2	77	5	ABP00402	ABP00402	Human ORF
118	20	95.2	56	5	ABG45247	ABG45247	Human pep		191	20	95.2	77	8	ABO55675	ABO55675	Human gen
119	20	95.2	56	6	ABM56447	ABM56447	Propionib		192	20	95.2	77	8	ADN99383	ADN99383	Novel hum
120	20	95.2	58	4	AAU04670	AAU04670	Human pol		193	20	95.2	78	4	ABG29189	ABG29189	Novel hum
121	20	95.2	58	5	ABP07646	ABP07646	Human ORF		194	20	95.2	78	5	ABP00882	ABP00882	Human ORF
122	20	95.2	58	7	ADG10562	ADG10562	Human STA		195	20	95.2	78	5	ABK34737	ABK34737	Novel hum
123	20	95.2	59	3	AAU01585	AAU01585	Human sec		196	20	95.2	79	5	ABP00882	ABP00882	Human ORF
124	20	95.2	59	4	AAU49510	AAU49510	Propionib		197	20	95.2	79	5	ABG29189	ABG29189	Novel hum
125	20	95.2	59	4	ABM46029	ABM46029	Propionib		198	20	95.2	80	5	ABP03115	ABP03115	Human ORF
126	20	95.2	60	5	ADK34339	ADK34339	Novel hum		199	20	95.2	80	5	ABG58738	ABG58738	Arabidops
127	20	95.2	61	3	AAU58739	AAU58739	Arabidops		200	20	95.2	80	5	ABG58738	ABG58738	Arabidops
128	20	95.2	61	3	AAU65071	AAU65071	Propionib		201	20	95.2	81	3	AAU91643	AAU91643	Human sec
129	20	95.2	61	4	AAU50814	AAU50814	Propionib		202	20	95.2	81	4	AAU84170	AAU84170	Human imm
130	20	95.2	61	6	ABM47333	ABM47333	Propionib		203	20	95.2	81	6	ABM56163	ABM56163	Propionib
131	20	95.2	61	6	ABM61590	ABM61590	Propionib		204	20	95.2	81	8	ADL1721	ADL1721	Novel hum
132	20	95.2	61	8	ABP76329	ABP76329	Human GEN		205	20	95.2	81	8	ADT57031	ADT57031	Plant pol
133	20	95.2	61	8	ABP96519	ABP96519	Novel S.		206	20	95.2	82	3	AAU91557	AAU91557	Human sec
134	20	95.2	61	9	ABM60389	ABM60389	Streptoco		207	20	95.2	82	3	AAU21352	AAU21352	Human nov
135	20	95.2	62	4	AAU47875	AAU47875	Human col		208	20	95.2	82	4	ADL71630	ADL71630	Novel hum
136	20	95.2	62	4	AAU59323	AAU59323	Propionib		209	20	95.2	82	4	AAU01170	AAU01170	Human pol
137	20	95.2	62	4	AAU59442	AAU59442	Propionib		210	20	95.2	83	4	AAU39795	AAU39795	Propionib
138	20	95.2	62	5	ABP02008	ABP02008	Human ORF		211	20	95.2	83	4	AAU61154	AAU61154	Propionib
139	20	95.2	62	6	ABM55961	ABM55961	Propionib		212	20	95.2	83	6	ABM57673	ABM57673	Propionib
140	20	95.2	62	6	ABM55842	ABM55842	Propionib		213	20	95.2	83	6	ABM36314	ABM36314	Propionib
141	20	95.2	63	4	AAU59231	AAU59231	Propionib		214	20	95.2	84	3	ABG20426	ABG20426	Arabidops
142	20	95.2	63	6	ABM55750	ABM55750	Propionib		215	20	95.2	85	3	ABM32591	ABM32591	Eucalyptu
143	20	95.2	63	6	ABU22642	ABU22642	Protein e		216	20	95.2	85	5	ABP00233	ABP00233	Human ORF
144	20	95.2	63	6	AAU21558	AAU21558	Protein e		217	20	95.2	85	5	AAE23312	AAE23312	Human p54
145	20	95.2	64	4	AAU48935	AAU48935	Propionib		218	20	95.2	85	5	AAE23318	AAE23318	Human p85
146	20	95.2	64	8	ABO57564	ABO57564	Human gen		219	20	95.2	85	7	ADD36316	ADD36316	D melanog
147	20	95.2	65	3	AAU32396	AAU32396	Propionib		220	20	95.2	85	7	ADD36330	ADD36330	D melanog
148	20	95.2	65	4	AAU54243	AAU54243	Propionib		221	20	95.2	85	8	ADQ08953	ADQ08953	D. melano
149	20	95.2	65	4	ABM17150	ABM17150	Human ner		222	20	95.2	85	8	ADQ08939	ADQ08939	D. melano
150	20	95.2	65	4	ABM17150	ABM17150	Human ner		223	20	95.2	85	9	ADW46118	ADW46118	Fruit fly
151	20	95.2	65	6	ABM50762	ABM50762	S. coelic		224	20	95.2	86	4	AAU46132	AAU46132	Fruit fly
152	20	95.2	65	6	ABM95052	ABM95052	Enterococ		225	20	95.2	86	4	AAU10340	AAU10340	Human pol
153	20	95.2	65	7	ADH87393	ADH87393	Human rep		226	20	95.2	86	4	AAU39146	AAU39146	Propionib
154	20	95.2	66	4	AAU94716	AAU94716	Human pol		227	20	95.2	86	4	AAU40456	AAU40456	Propionib
155	20	95.2	66	4	AAU01640	AAU01640	Human pol		228	20	95.2	86	6	ABM35665	ABM35665	Propionib
156	20	95.2	66	4	AAU52583	AAU52583	Propionib		229	20	95.2	86	6	ABM36975	ABM36975	Propionib
157	20	95.2	66	4	AAU22743	AAU22743	Human pro		230	20	95.2	86	6	ABU70552	ABU70552	Human adi
158	20	95.2	66	6	ABM49102	ABM49102	Propionib		231	20	95.2	86	9	ADZ10664	ADZ10664	N. europe
159	20	95.2	66	7	ADU09316	ADU09316	Human pro		232	20	95.2	86	9	AAW35568	AAW35568	Borna dis
160	20	95.2	67	4	AAU21127	AAU21127	Human nov		233	20	95.2	87	2	AAW35568	AAW35568	Borna dis
161	20	95.2	68	4	AAU92836	AAU92836	Human dig		234	20	95.2	87	5	ABY79287	ABY79287	Amino aci
162	20	95.2	68	4	AAU90490	AAU90490	Human imm		235	20	95.2	87	7	ADD36276	ADD36276	D melanog
163	20	95.2	68	4	ABM11674	ABM11674	Human rib		236	20	95.2	87	8	ADL05617	ADL05617	M. catarr
164	20	95.2	68	4	AAU80273	AAU80273	Human pro		237	20	95.2	87	8	ADQ08899	ADQ08899	Anopheles
165	20	95.2	68	4	AAU80273	AAU80273	Human pro		238	20	95.2	87	9	ADW46078	ADW46078	Fruit fly
166	20	95.2	69	5	ABP34005	ABP34005	Human nov		239	20	95.2	88	5	ABP28776	ABP28776	Streptoco
167	20	95.2	70	4	AAU51980	AAU51980	Propionib		240	20	95.2	88	8	ADV89222	ADV89222	Streptoco
168	20	95.2	70	6	ABM48499	ABM48499	Propionib		241	20	95.2	88	8	ADV80475	ADV80475	Streptoco
169	20	95.2	71	2	AAU76577	AAU76577	Human ova		242	20	95.2	88	8	ADV82536	ADV82536	Streptoco
170	20	95.2	71	4	AAU54982	AAU54982	Propionib		243	20	95.2	89	4	AAU45557	AAU45557	Propionib

244	20	95.2	89	6	ABM42076	Abm42076 Propionib	317	20	95.2	111	5	ABP41090	Abp41090 Human ova
245	20	95.2	90	3	AAB13265	Aab13265 Caenrhab	318	20	95.2	111	7	ABM86150	Abm86150 Rice abio
246	20	95.2	90	4	AAB78982	Aab78982 C. glutam	319	20	95.2	111	8	ADL81889	Adl81889 P. aerugi
247	20	95.2	90	5	ABB78168	Abb78168 Amino aci	320	20	95.2	112	4	AAE01912	Aae01912 Arabidops
248	20	95.2	90	8	ADX78594	Adx78594 Plant ful	321	20	95.2	112	4	AAE02467	Aae02467 Arabidops
249	20	95.2	91	4	AAG75213	Aag75213 Human col	322	20	95.2	112	7	ADP55784	Adp55784 Thalecres
250	20	95.2	91	6	ABU70576	Abu70576 Human adi	323	20	95.2	112	7	ADD30554	Add30554 Plant yie
251	20	95.2	92	4	AAW25520	Aaw25520 Human pro	324	20	95.2	112	7	ADP37115	Adp37115 Plant yie
252	20	95.2	93	7	ADC94062	Adc94062 E. faeciu	325	20	95.2	112	8	ADJ41575	Adj41575 Plant tra
253	20	95.2	94	4	ABB74198	Abb74198 Human col	326	20	95.2	112	8	ADJ56850	Adj56850 Arabidops
254	20	95.2	94	4	AAB70492	Aab70492 Drosophil	327	20	95.2	112	8	ADO03337	Ado03337 Thalecres
255	20	95.2	94	4	AAD60007	Aad60007 Propionib	328	20	95.2	112	8	ADO03603	Ado03603 Thalecres
256	20	95.2	94	4	ABG11744	Abg11744 Novel hum	329	20	95.2	112	8	ADO01729	Ado01729 Thalecres
257	20	95.2	94	5	ABP25808	Abp25808 Streptoco	330	20	95.2	112	8	ADO62911	Ado62911 Transcrip
258	20	95.2	94	6	ABM56526	Abm56526 Propionib	331	20	95.2	112	8	ADS16859	Ads16859 Thale cre
259	20	95.2	94	8	ADV89513	Adv89513 Streptoco	332	20	95.2	112	9	AEA26445	Aea26445 Stres to
260	20	95.2	94	8	ADV82929	Adv82929 Streptoco	333	20	95.2	113	4	AAO09987	Aao09987 Human pol
261	20	95.2	94	8	ADV80766	Adv80766 Streptoco	334	20	95.2	113	6	ADB10752	Adb10752 Allolococ
262	20	95.2	95	3	AAV79286	Aav79286 Borna dis	335	20	95.2	113	7	ADM25417	Adm25417 Hyperther
263	20	95.2	95	3	AAG51932	Aag51932 Arabidops	336	20	95.2	114	4	AGG91809	Agg91809 C glutami
264	20	95.2	95	3	AAO20632	Aao20632 Arabidops	337	20	95.2	115	3	AAV91642	Aav91642 Human sec
265	20	95.2	95	4	ABB11442	Abb11442 Human sec	338	20	95.2	115	4	AAV91827	Aav91827 C glutami
266	20	95.2	95	4	AAU49110	Aau49110 Propionib	339	20	95.2	115	5	ABP08923	Abp08923 Human ORP
267	20	95.2	95	4	ABG18078	Abg18078 Novel hum	340	20	95.2	115	6	ABU70451	Abu70451 Human adi
268	20	95.2	95	6	ABM45629	Abm45629 Propionib	341	20	95.2	115	8	ADL71720	Adl71720 Novel hum
269	20	95.2	95	8	ADX93779	Adx93779 Plant ful	342	20	95.2	115	8	ADP32944	Adp32944 Human pho
270	20	95.2	95	8	ADY24463	Ady24463 Plant ful	343	20	95.2	116	5	AAH82867	Aah82867 Human imm
271	20	95.2	96	3	AAZ29476	Aaz29476 Burkholde	344	20	95.2	117	4	AAH82867	Aah82867 Human imm
272	20	95.2	96	4	ABG21787	Abg21787 Novel hum	345	20	95.2	117	4	ABP5436	Abp5436 Human pro
273	20	95.2	97	2	AAV37141	Aav37141 Amino aci	346	20	95.2	118	5	ABP27226	Abp27226 Streptoco
274	20	95.2	97	4	AAU55876	Aau55876 Propionib	347	20	95.2	118	6	ABU46798	Abu46798 Protein e
275	20	95.2	97	6	AAU52395	Aau52395 Propionib	348	20	95.2	118	6	ABU76051	Abu76051 Human GEN
276	20	95.2	97	7	ABO79693	Abu79693 Pseudomon	349	20	95.2	118	8	ADY75855	Ady75855 Plant ful
277	20	95.2	98	3	AAO34149	Aao34149 Zea mays	350	20	95.2	119	3	AAV75271	Aav75271 Neisseria
278	20	95.2	98	4	ABB63383	Abb63383 Drosophil	351	20	95.2	119	5	ABP40144	Abp40144 Staphyloc
279	20	95.2	98	4	AAU47172	Aau47172 Propionib	352	20	95.2	119	7	ADO77643	Ado77643 Pseudomon
280	20	95.2	98	5	ABP07670	Abp07670 Human ORP	353	20	95.2	119	8	ADO77188	Ado77188 Staphyloc
281	20	95.2	98	6	ABM43691	Abm43691 Propionib	354	20	95.2	120	3	AAO49392	Aao49392 Arabidops
282	20	95.2	99	5	ABR40491	Abm40491 Humah sec	355	20	95.2	120	4	AAH85739	Aah85739 Human imm
283	20	95.2	99	5	ABR40417	Abm40417 Human sec	356	20	95.2	120	4	ABG12278	Abg12278 Human sec
284	20	95.2	99	5	ABP31616	Abp31616 Human ORP	357	20	95.2	120	4	ABG11980	Abg11980 Novel hum
285	20	95.2	100	3	AAO01140	Aao01140 Humah sec	358	20	95.2	121	2	AAW77543	Aaw77543 Staphyloc
286	20	95.2	100	3	AAE55172	Aae55172 Arabidops	359	20	95.2	121	3	AAO49391	Aao49391 Arabidops
287	20	95.2	100	4	ABB66078	Abb66078 Drosophil	360	20	95.2	121	4	AAU48378	Aau48378 Propionib
288	20	95.2	100	4	AAU91091	Aau91091 Humah imm	361	20	95.2	121	4	AAU44395	Aau44395 Propionib
289	20	95.2	100	4	ABG16166	Abg16166 Novel hum	362	20	95.2	121	6	ABM44897	Abm44897 Propionib
290	20	95.2	100	8	ADJ57835	Adj57835 WRB prote	363	20	95.2	121	6	ABM40914	Abm40914 Propionib
291	20	95.2	101	6	ADB11138	Adb11138 Allolococ	364	20	95.2	122	3	AAO37576	Aao37576 Arabidops
292	20	95.2	102	4	AAU50064	Aau50064 Propionib	365	20	95.2	122	3	AAO49390	Aao49390 Arabidops
293	20	95.2	102	6	ABM46583	Abm46583 Propionib	366	20	95.2	122	5	ABG64896	Abg64896 Human alb
294	20	95.2	104	4	AAO13013	Aao13013 Humah pol	367	20	95.2	122	5	ABR77034	Abb77034 Human pro
295	20	95.2	104	4	AAO82569	Aag82569 S. epider	368	20	95.2	122	7	ADC88716	Adc88716 Ribosomal
296	20	95.2	105	4	AAW24351	Aam24351 Human EST	369	20	95.2	122	7	ABO74609	Abu74609 Pseudomon
297	20	95.2	105	7	ADC31355	Adc31355 Humah nov	370	20	95.2	122	8	ADL78163	Adl78163 Albumin f
298	20	95.2	106	3	AAO05441	Aag05441 Arabidops	371	20	95.2	123	3	AAO39327	Aao39327 Human sec
299	20	95.2	106	4	ABB66482	Abb66482 Drosophil	372	20	95.2	123	5	ABP64793	Abp64793 Human pro
300	20	95.2	106	4	ABM66481	Abb66481 Propionib	373	20	95.2	123	7	ADM04917	Adm04917 Human pro
301	20	95.2	106	4	AAU14407	Aau14407 Humah nov	374	20	95.2	123	7	ABM87812	Abm87812 Rice abio
302	20	95.2	106	4	AAU14408	Aau14408 Human nov	375	20	95.2	123	8	ABM81082	Abm81082 Tumour-as
303	20	95.2	106	7	ADE09443	Adh09443 Novel pro	376	20	95.2	123	9	ADX08058	Adx08058 Cyclin-de
304	20	95.2	106	8	ADH80726	Adh80726 Humah pol	377	20	95.2	124	7	ABM74443	Abm74443 CNA-clone
305	20	95.2	106	8	ADH80725	Adh80725 Humah pol	378	20	95.2	125	8	AAV85832	Aav85832 S. pneumo
306	20	95.2	107	4	AAU67163	Aau67163 Propionib	379	20	95.2	125	8	ADK46925	Adk46925 Streptoco
307	20	95.2	107	5	ABP68823	Abp68823 Humah pol	380	20	95.2	126	3	AAO32392	Aao32392 Arabidops
308	20	95.2	107	5	ABP38042	Abp38042 Staphyloc	381	20	95.2	126	5	ABP33115	Abp33115 Human red
309	20	95.2	107	6	ABM63682	Abm63682 Propionib	382	20	95.2	126	6	ABR40675	Abm40675 Zea mays
310	20	95.2	109	4	AAU37519	Aau37519 Staphyloc	383	20	95.2	126	8	ABM80250	Abm80250 Tumour-as
311	20	95.2	109	8	ADO66927	Adq66927 Novel hum	384	20	95.2	127	4	AAU64816	Aau64816 Propionib
312	20	95.2	110	4	AAU45128	Aau45128 Propionib	385	20	95.2	127	6	ABM61335	Abm61335 Propionib
313	20	95.2	110	4	AAU32624	Aau32624 Novel hum	386	20	95.2	127	8	ADM26194	Adm26194 Hyperther
314	20	95.2	110	6	ABM41647	Abm41647 Propionib	387	20	95.2	127	8	ADX66219	Adx66219 Plant ful
315	20	95.2	111	3	AAH43863	Aah43863 Human can	388	20	95.2	128	4	AAO98867	Aao98867 E. coli g
316	20	95.2	111	3	AAB56865	Aab56865 Human pro	389	20	95.2	128	5	ADK36004	Adk36004 Novel hum

390	20	95.2	129	5	ABP34766	Abp34766 Human iso	463	20	95.2	149	4	AAU19479	Aau19479 Human dia
391	20	95.2	129	8	ABM80361	Abm80361 Tumour-as	464	20	95.2	149	6	ABM69638	Abm69638 Photorhab
392	20	95.2	129	9	ABM94297	Abm94297 M. xanthu	465	20	95.2	149	7	ADF59714	Adf59714 Human pol
393	20	95.2	130	4	ABU53192	Abu53192 Human tes	466	20	95.2	149	7	ADO83188	Ado83188 Pseudomon
394	20	95.2	130	8	ADN05496	Adn05496 Antipseri	467	20	95.2	149	8	ADI67173	Adi67173 Lactobaci
395	20	95.2	131	3	AGS56232	Agss56232 Arabidops	468	20	95.2	149	8	ADR94703	Adr94703 Novel S.
396	20	95.2	131	3	AGS56232	Agss56232 Arabidops	469	20	95.2	149	8	ADU733960	Adu733960 Melampor
397	20	95.2	131	4	AAU23240	Aau23240 Novel hum	470	20	95.2	149	9	AEA58573	Aea58573 Streptoco
398	20	95.2	132	8	ADU79464	Adu79464 Plant ful	471	20	95.2	150	3	AGS58826	Agss58826 Arabidops
399	20	95.2	132	8	ADU79474	Adu79474 Plant ful	472	20	95.2	150	5	ABU11280	Abu11280 Yeast sel
400	20	95.2	133	4	ABM70513	Abm70513 Drosophil	473	20	95.2	150	7	ADU60917	Adu60917 Human pro
401	20	95.2	133	7	ABO61598	Abom61598 Klebsiell	474	20	95.2	150	7	ADJ68771	Adj68771 Human hea
402	20	95.2	134	2	AAU32292	Aau32292 Sequence	475	20	95.2	150	7	ADU73953	Adu73953 Melampor
403	20	95.2	134	3	AAU08210	Aau08210 Arabidops	476	20	95.2	150	8	ADU73955	Adu73955 Melampor
404	20	95.2	135	3	AAU03345	Aau03345 Human sec	477	20	95.2	150	8	ADU73959	Adu73959 Melampor
405	20	95.2	135	3	AAU26193	Aau26193 Arabidops	478	20	95.2	150	8	AAU23136	Aau23136 Arabidops
406	20	95.2	135	8	ADU73541	Adu73541 Plant ful	479	20	95.2	151	3	AAU59190	Aau59190 Arabidops
407	20	95.2	136	4	AAU057372	Aau057372 Propionib	480	20	95.2	151	4	AAU56629	Aau56629 Propionib
408	20	95.2	136	4	AAU04765	Aau04765 Human ves	481	20	95.2	151	4	AAU56629	Aau56629 Propionib
409	20	95.2	136	6	ABM53891	Abm53891 Propionib	482	20	95.2	151	6	ABM578723	Abm578723 N. gonorr
410	20	95.2	136	6	ABU70404	Abu70404 Human adi	483	20	95.2	151	6	ABM53148	Abm53148 Propionib
411	20	95.2	136	8	ABM80151	Abm80151 Tumour-as	484	20	95.2	151	6	ABU38015	Abu38015 Protein e
412	20	95.2	137	3	AAU08209	Aau08209 Arabidops	485	20	95.2	151	6	ABU37100	Abu37100 Protein e
413	20	95.2	137	6	ABU070769	Abu070769 Human adi	486	20	95.2	151	8	ADU09612	Adu09612 Plant ful
414	20	95.2	139	4	ABU43303	Abu43303 Peptide #	487	20	95.2	152	5	ABP26065	Abp26065 Streptoco
415	20	95.2	139	4	AAU37148	Aau37148 Human bon	488	20	95.2	152	6	ABM15891	Abm15891 Mycobacte
416	20	95.2	139	4	AAU64211	Aau64211 Human bra	489	20	95.2	152	6	AAU57110	Aau57110 Human pro
417	20	95.2	139	4	ABU58696	Abu58696 Human liv	490	20	95.2	153	3	AAU40433	Aau40433 Arabidops
418	20	95.2	139	7	ABU076363	Abu076363 Pseudomon	491	20	95.2	154	3	AAU40433	Aau40433 Arabidops
419	20	95.2	139	7	ABU076368	Abu076368 Pseudomon	492	20	95.2	154	5	ABP64628	Abp64628 Human ORF
420	20	95.2	139	9	ABU076368	Abu076368 Pseudomon	493	20	95.2	154	8	ADU92363	Adu92363 Plant ful
421	20	95.2	140	5	ABG77378	Abg77378 Selected	494	20	95.2	155	4	AAU56468	Aau56468 Propionib
422	20	95.2	140	7	ADU00833	Adu00833 Enterohae	495	20	95.2	155	6	ABM52987	Abm52987 Propionib
423	20	95.2	140	8	ADU06007	Adu06007 hMash-1-d	496	20	95.2	155	9	ABM5565	Abm5565 M. xanthu
424	20	95.2	140	8	ADU27672	Adu27672 Human ach	497	20	95.2	156	4	AAU41931	Aau41931 Human pol
425	20	95.2	140	9	ADU27672	Adu27672 Human ach	498	20	95.2	157	4	ABM70572	Abm70572 Drosophil
426	20	95.2	141	6	ABU40239	Abu40239 Protein e	499	20	95.2	157	6	ABP79900	Abp79900 N. gonorr
427	20	95.2	142	3	ABU59091	Abu59091 Breat an	500	20	95.2	158	3	AAU21092	Aau21092 Arabidops
428	20	95.2	142	4	ABG27260	Abg27260 Novel hum	501	20	95.2	158	3	AAU05681	Aau05681 Arabidops
429	20	95.2	142	7	ABG69587	Abg69587 Pseudomon	502	20	95.2	158	3	AAU25206	Aau25206 Arabidops
430	20	95.2	143	5	AAU17807	Aau17807 Herpes si	503	20	95.2	158	4	AAU49514	Aau49514 Propionib
431	20	95.2	143	7	ADG74939	Adg74939 Human her	504	20	95.2	158	6	ABM46033	Abm46033 Propionib
432	20	95.2	144	4	AAU36192	Aau36192 Pseudomon	505	20	95.2	158	6	ABM46033	Abm46033 Propionib
433	20	95.2	144	5	ABP29247	Abp29247 Streptoco	506	20	95.2	158	8	ADU13492	Adu13492 Linking r
434	20	95.2	144	6	ABU38249	Abu38249 Protein e	507	20	95.2	158	8	ADK47131	Adk47131 Streptoco
435	20	95.2	144	7	ADU089174	Adu089174 Ribosomal	508	20	95.2	158	8	ADU95100	Adu95100 Novel S.
436	20	95.2	144	7	ADU089161	Adu089161 Ribosomal	509	20	95.2	158	8	ADU89708	Adu89708 Streptoco
437	20	95.2	144	7	ADU60089	Adu60089 Rat Prote	510	20	95.2	158	8	ADU80961	Adu80961 Streptoco
438	20	95.2	144	7	ADU60091	Adu60091 Human Pro	511	20	95.2	158	8	ADU88039	Adu88039 Plant ful
439	20	95.2	144	7	ADU45769	Adu45769 Rat Prote	512	20	95.2	158	8	ADU88039	Adu88039 Plant ful
440	20	95.2	144	7	ADU60093	Adu60093 Rat Prote	513	20	95.2	159	2	AAU60121	Aau60121 Human end
441	20	95.2	144	7	ADU60095	Adu60095 Human Pro	514	20	95.2	159	2	AAU60121	Aau60121 Human end
442	20	95.2	144	7	ADU60095	Adu60095 Human Pro	515	20	95.2	159	6	ABM60328	Abm60328 Propionib
443	20	95.2	144	7	ADU45771	Adu45771 Human Pro	516	20	95.2	160	5	AAU15947	Aau15947 Fly CP334
444	20	95.2	145	1	AAU81003	Aau81003 Sequence	517	20	95.2	160	5	AAU15946	Aau15946 Fly CP334
445	20	95.2	145	7	ADU85279	Adu85279 Rat DNA b	518	20	95.2	160	5	AAU15946	Aau15946 Fly CP334
446	20	95.2	145	7	ADU89175	Adu89175 Ribosomal	519	20	95.2	160	5	AAU15946	Aau15946 Fly CP334
447	20	95.2	145	7	ADU65606	Adu65606 Rat Prote	520	20	95.2	160	5	AAU15946	Aau15946 Fly CP334
448	20	95.2	145	8	ABM81988	Abm81988 Tumour-as	521	20	95.2	161	2	AAU48359	Aau48359 Human pro
449	20	95.2	145	8	ADU60158	Adu60158 Housekeep	522	20	95.2	161	2	AAU48359	Aau48359 Human pro
450	20	95.2	145	8	ADU60158	Adu60158 Housekeep	523	20	95.2	162	3	AAU48359	Aau48359 Human pro
451	20	95.2	146	4	AAU00377	Aau00377 Human pol	524	20	95.2	162	7	ADU40432	Adu40432 Arabidops
452	20	95.2	146	7	ADU60911	Adu60911 Rat Prote	525	20	95.2	162	8	ADK47792	Adk47792 Streptoco
453	20	95.2	146	7	ADU60915	Adu60915 Rat Prote	526	20	95.2	162	8	ADK47792	Adk47792 Streptoco
454	20	95.2	146	8	ADU90949	Adu90949 Plant ful	527	20	95.2	166	4	ABM62673	Abm62673 Streptoco
455	20	95.2	146	8	ADU90949	Adu90949 Plant ful	528	20	95.2	166	4	ABM62673	Abm62673 Streptoco
456	20	95.2	147	4	ABG24035	Abg24035 Plant ful	529	20	95.2	167	4	ABM62673	Abm62673 Streptoco
457	20	95.2	147	5	ABU97880	Abu97880 Human sec	530	20	95.2	167	4	ABM62673	Abm62673 Streptoco
458	20	95.2	147	5	ABU99297	Aau99297 Pyrococcu	531	20	95.2	167	4	ABM62673	Abm62673 Streptoco
459	20	95.2	147	6	ABU02446	Abu02446 S. pneumo	532	20	95.2	167	4	ABM62673	Abm62673 Streptoco
460	20	95.2	147	8	ADU15865	Adu15865 Predicted	533	20	95.2	168	7	ADU12762	Adu12762 Plant ful
461	20	95.2	147	8	ADU15865	Adu15865 Predicted	534	20	95.2	168	8	ADU12762	Adu12762 Plant ful
462	20	95.2	149	3	AAU41916	Aau41916 Human ORF	535	20	95.2	168	9	AAU41916	Aau41916 Human ORF



536	20	95.2	169	7	ABO72175	Pseudomon	609	20	95.2	192	3	AAG28658	Aag28658 Arabidops
537	20	95.2	169	8	ADN26043	Bacterial	610	20	95.2	192	3	AAG53689	Aag53689 Arabidops
538	20	95.2	170	3	ABE25298	Eucalyptu	611	20	95.2	192	4	AAU45522	AAu45522 Propionib
539	20	95.2	170	4	AAU62022	Propionib	612	20	95.2	192	4	ABG23080	Abg23080 Novel hum
540	20	95.2	170	6	ABM58541	Propionib	613	20	95.2	192	6	ABM42041	ABm42041 Propionib
541	20	95.2	170	7	ABO74371	Pseudomon	614	20	95.2	192	8	ADR20249	Adr20249 Herv-K/HM
542	20	95.2	170	7	ABO69833	Pseudomon	615	20	95.2	192	8	ADR95795	Adr95795 Novel S.
543	20	95.2	170	8	ADR94851	Novel S.	616	20	95.2	192	8	ADT56323	Adt56323 Plant pol
544	20	95.2	170	9	AEA58721	Streptoco	617	20	95.2	192	9	ADW72805	Adw72805 Human Nkp
545	20	95.2	172	7	ABO73955	Pseudomon	618	20	95.2	192	9	AEA40203	Aea40203 Human nat
546	20	95.2	172	8	ADY12741	Plant ful	619	20	95.2	192	9	AEA59665	Aea59665 Streptoco
547	20	95.2	172	9	ABM90696	M. xanthu	620	20	95.2	193	4	ABE66326	ABe66326 Drosophil
548	20	95.2	172	9	ABM93767	M. xanthu	621	20	95.2	193	7	ABO75621	ABo75621 Pseudomon
549	20	95.2	173	4	ABG11981	Novel hum	622	20	95.2	193	9	ABM93373	ABm93373 M. xanthu
550	20	95.2	173	6	ADA33948	Acinetoba	623	20	95.2	194	3	ABG08208	ABg08208 Arabidops
551	20	95.2	173	8	ADR22520	Human tra	624	20	95.2	194	5	ABO80622	ABo80622 Human DNA
552	20	95.2	174	5	AAU78275	Humah rap	625	20	95.2	195	8	ADT57977	Adt57977 Plant pol
553	20	95.2	175	4	AAU41718	Humah pol	626	20	95.2	196	4	ABG11745	ABg11745 Novel hum
554	20	95.2	175	4	AAU27520	Humah G-P	627	20	95.2	196	4	ABG11745	ABg11745 Novel hum
555	20	95.2	175	5	AAU99296	Pyrotoocu	628	20	95.2	196	6	ABR41472	ABr41472 Human DIT
556	20	95.2	175	8	ADT60505	Plant pol	629	20	95.2	197	4	ABG16546	ABg16546 Novel hum
557	20	95.2	176	8	ADT57037	Plant pol	630	20	95.2	198	3	ABG01347	ABg01347 Human sec
558	20	95.2	176	6	ABX70906	Plant ful	631	20	95.2	199	7	ADM26103	Adm26103 Hyperther
559	20	95.2	177	6	ABU23025	Protein e	632	20	95.2	199	7	ABO74256	ABo74256 Pseudomon
560	20	95.2	178	7	ABO79964	Pseudomon	633	20	95.2	199	8	ADU17770	Adu17770 Reverse t
561	20	95.2	179	3	AGS32991	Arabidops	634	20	95.2	199	8	ADU17793	Adu17793 Reverse t
562	20	95.2	179	8	ADY83107	Streptoco	635	20	95.2	199	8	ADU17783	Adu17783 Reverse t
563	20	95.2	180	3	ABE24670	Plant SDF	636	20	95.2	199	8	ADU17794	Adu17794 Reverse t
564	20	95.2	180	3	ABE24658	Plant SDF	637	20	95.2	199	8	ADU17767	Adu17767 Reverse t
565	20	95.2	180	3	ABE59347	Arabidops	638	20	95.2	199	8	ADU17768	Adu17768 Reverse t
566	20	95.2	180	7	ADC87059	Humah GPC	639	20	95.2	199	8	ADU17766	Adu17766 Reverse t
567	20	95.2	180	7	ADP07384	Bacterial	640	20	95.2	199	8	ADU17772	Adu17772 Reverse t
568	20	95.2	180	8	ADU17867	Reverse t	641	20	95.2	199	8	ADU17795	Adu17795 Reverse t
569	20	95.2	180	8	ADU17864	Reverse t	642	20	95.2	200	2	AAW28317	AAw28317 Amino aci
570	20	95.2	180	9	ADY27674	Humah ach	643	20	95.2	201	3	ABG21619	ABg21619 Arabidops
571	20	95.2	182	4	ABG16757	Novel hum	644	20	95.2	201	4	ABE63738	ABe63738 Human pro
572	20	95.2	182	7	ABO71799	Pseudomon	645	20	95.2	201	7	ABO79454	ABo79454 Pseudomon
573	20	95.2	183	8	ADL05999	M. catarr	646	20	95.2	202	5	ABE90006	ABe90006 Human pol
574	20	95.2	183	8	ADX67620	Plant ful	647	20	95.2	202	7	ABO73570	ABo73570 Pseudomon
575	20	95.2	183	8	ADX78430	Plant ful	648	20	95.2	203	4	ABE93444	ABe93444 Human pro
576	20	95.2	184	4	ABE64436	Humah sec	649	20	95.2	203	5	ABE89562	ABe89562 Human pol
577	20	95.2	185	4	ABG19120	Novel hum	650	20	95.2	203	6	ABU56465	ABu56465 Lung canc
578	20	95.2	185	7	ADM25943	Hyperther	651	20	95.2	203	8	ADH18880	ADh18880 Human cel
579	20	95.2	185	7	ABO62926	Klebsiell	652	20	95.2	203	8	ADN05076	ADn05076 Antipsori
580	20	95.2	185	8	ADL06062	M. catarr	653	20	95.2	205	4	AAU67257	AAu67257 Propionib
581	20	95.2	185	8	ADX88139	Plant ful	654	20	95.2	205	5	ABE94341	ABe94341 Chlamydia
582	20	95.2	187	2	AAV38513	Neisseria	655	20	95.2	205	5	ABU05791	ABu05791 M. tuberc
583	20	95.2	187	3	ABM03148	Humah neu	656	20	95.2	205	5	ABU05386	ABu05386 M. tuberc
584	20	95.2	187	9	ABM93390	M. xanthu	657	20	95.2	205	6	ABM63776	ABm63776 Propionib
585	20	95.2	187	9	ABM91093	M. xanthu	658	20	95.2	205	7	ABO76651	ABo76651 Pseudomon
586	20	95.2	187	9	ABE48883	N. mening	659	20	95.2	206	2	AAW81881	AAw81881 Mouse typ
587	20	95.2	188	7	ADP07023	Bacterial	660	20	95.2	206	2	AAW48977	AAw48977 Mouse OX4
588	20	95.2	188	7	ADM25787	Hyperther	661	20	95.2	207	8	ADJ48873	ADj48873 Oil-assoc
589	20	95.2	188	8	ABO59539	Humah gen	662	20	95.2	208	5	ABP38028	ABp38028 Staphyloc
590	20	95.2	188	8	ADY07159	Plant ful	663	20	95.2	208	6	ABE52585	ABe52585 Protein s
591	20	95.2	189	3	AAE43607	Humah can	664	20	95.2	208	7	ADK63902	ADk63902 Disease t
592	20	95.2	189	3	AAE60392	Arabidops	665	20	95.2	208	7	ABO82964	ABo82964 Pseudomon
593	20	95.2	189	3	AAE57484	Arabidops	666	20	95.2	209	4	ABE86656	ABe86656 Drosophil
594	20	95.2	189	7	ADD18799	Humah dis	667	20	95.2	209	4	ABG15663	ABg15663 Novel hum
595	20	95.2	189	8	ABO59519	Humah gen	668	20	95.2	209	9	ADW72804	Adw72804 Human Nkp
596	20	95.2	189	8	ABO58870	Humah pro	669	20	95.2	209	9	AEA40202	Aea40202 Human nat
597	20	95.2	189	9	ADY18038	PRO Polyp	670	20	95.2	210	4	AAE63739	AAe63739 Human pro
598	20	95.2	190	3	AAE32452	Arabidops	671	20	95.2	210	4	AAE63742	AAe63742 Human pro
599	20	95.2	190	3	AAE26346	Arabidops	672	20	95.2	210	5	ABU05448	ABu05448 M. tuberc
600	20	95.2	190	4	ABE70193	Drosophil	673	20	95.2	210	6	ABU41414	ABu41414 Protein e
601	20	95.2	190	7	ABO80877	Pseudomon	674	20	95.2	210	8	ADT57410	ADt57410 Plant pol
602	20	95.2	191	4	ABE66296	Drosophil	675	20	95.2	211	8	ADX97087	ADx97087 Plant ful
603	20	95.2	191	8	ADY09572	Plant ful	676	20	95.2	211	4	AAE18601	AAe18601 Peptide #
604	20	95.2	192	2	AAW64369	Mycobacte	677	20	95.2	211	4	ABE37657	ABe37657 Peptide #
605	20	95.2	192	2	AAW81734	M. tuberc	678	20	95.2	211	4	AAE31058	AAe31058 Peptide #
606	20	95.2	192	2	AAV39023	M. tuberc	679	20	95.2	211	4	ABE32378	ABe32378 Peptide #
607	20	95.2	192	2	AAV39166	M. tuberc	680	20	95.2	211	4	AAE39932	AAe39932 Human pol
608	20	95.2	192	3	AAE26192	Arabidops	681	20	95.2	211	4	AAE58289	AAe58289 Human bra

682	20	95.2	211	4	ABG52465	Abg52465 Human liv	755	20	95.2	231	5	ABB94106	Abb94106 Chlamydia
683	20	95.2	211	4	AAm06172	Aam06172 Peptide #	756	20	95.2	231	5	AAE18736	Aae18736 Human lun
684	20	95.2	211	5	ABg40501	Abg40501 Human pep	757	20	95.2	231	6	ABM71291	Abm71291 Staphyloc
685	20	95.2	211	8	ADx91661	Adx91661 Plant ful	758	20	95.2	231	8	ADY05007	Ady05007 Plant ful
686	20	95.2	212	2	AAy35210	Aay35210 Chlamydia	759	20	95.2	232	8	ADK70536	Adk70536 Respirato
687	20	95.2	212	3	AAg21618	Aag21618 Arabidops	760	20	95.2	233	3	AAy74686	Aay74686 Neisseria
688	20	95.2	212	3	AAU19495	Aau19495 Human dia	761	20	95.2	233	3	AAy74686	Aay74686 Neisseria
689	20	95.2	212	4	AAU67171	Aau67171 Propionib	762	20	95.2	233	4	ABG24475	Abg24475 Novel hum
690	20	95.2	212	4	ABG05041	Abg05041 Novel hum	763	20	95.2	233	4	ABG29906	Abg29906 Novel hum
691	20	95.2	212	4	ABM63690	Abm63690 Propionib	764	20	95.2	233	4	ABP27713	Abp27713 Streptoco
692	20	95.2	212	7	ABO77775	Abu77775 Pseudomon	765	20	95.2	233	5	AAE18735	Aae18735 Human lun
693	20	95.2	213	4	AAU19186	Aau19186 Human G P	766	20	95.2	233	5	AAE18735	Aae18735 Human lun
694	20	95.2	213	4	AAU50790	Aau50790 Propionib	767	20	95.2	233	5	ABP75496	Abp75496 Human sec
695	20	95.2	213	6	ABU02007	Abu02007 S. pneumo	768	20	95.2	233	6	ABU06043	Abu06043 N. lactam
696	20	95.2	213	6	ABM47309	Abm47309 Propionib	769	20	95.2	233	6	ADb64902	Adb64902 Human pro
697	20	95.2	213	7	ADP59061	Adp59061 Human pol	770	20	95.2	233	7	ADf59565	Adf59565 Human pol
698	20	95.2	213	7	ABO82012	Abu82012 Pseudomon	771	20	95.2	234	4	AAy59835	Aay59835 Human nor
699	20	95.2	214	4	AAU47428	Aau47428 Propionib	772	20	95.2	234	4	ABg21806	Abg21806 Novel hum
700	20	95.2	214	4	ABG00921	Abg00921 Novel hum	773	20	95.2	234	4	AAU85594	Aau85594 Lung tumo
701	20	95.2	214	6	AAE32354	Aae32354 Murine FD	774	20	95.2	234	5	ABU85592	Abu85592 Lung tumo
702	20	95.2	214	6	AAE43347	Aae43347 Propionib	775	20	95.2	234	5	ABU85592	Abu85592 Lung tumo
703	20	95.2	214	7	ADD25201	Add25201 Fertility	776	20	95.2	234	6	ABP81497	Abp81497 Lung can
704	20	95.2	214	7	ADD69285	Add69285 Murine FD	777	20	95.2	234	6	ABU66469	Abu66469 Lung can
705	20	95.2	214	8	ADN61216	Adn61216 Radish nu	778	20	95.2	234	7	ADH47391	Adh47391 Human lun
706	20	95.2	215	5	ABP30487	Abp30487 Streptoco	779	20	95.2	234	7	ADK47790	Adk47790 Streptoco
707	20	95.2	215	8	ADU69658	Adu69658 S. agalact	780	20	95.2	234	8	ADJ21310	Adj21310 Human lun
708	20	95.2	215	8	ADV88193	Adv88193 Streptoco	781	20	95.2	234	8	ADJ21310	Adj21310 Human lun
709	20	95.2	215	8	ADV79446	Adv79446 Streptoco	782	20	95.2	234	9	AEb03554	Aeb03554 Mycobacte
710	20	95.2	215	8	ADU81621	Adu81621 Human sig	783	20	95.2	234	9	AEa79457	Aea79457 Novel M.
711	20	95.2	216	4	ABU52791	Abu52791 Human rubb	784	20	95.2	235	2	AAr70101	Aar70101 LiYV RNA
712	20	95.2	216	5	ABG79610	Abg79610 Plant ful	785	20	95.2	235	3	AAg19255	Aag19255 Arabidops
713	20	95.2	216	5	ADV09061	Adv09061 Plant ful	786	20	95.2	235	3	AAg43087	Aag43087 Arabidops
714	20	95.2	217	5	AAE13525	Aae13525 Mouse FDR	787	20	95.2	235	3	AAU46627	Aau46627 Propionib
715	20	95.2	217	5	AAE13525	Aae13525 Mouse FDR	788	20	95.2	235	3	AAU46627	Aau46627 Propionib
716	20	95.2	218	3	AAU23415	Aau23415 Novel hum	789	20	95.2	236	2	AAy38516	Aay38516 Neisseria
717	20	95.2	219	4	ABG60165	Abg60165 Drosophil	790	20	95.2	236	2	ABG22704	Abg22704 Novel hum
718	20	95.2	220	3	ABG08911	Abg08911 Streptomy	791	20	95.2	236	2	AAU85591	Aau85591 Lung tumo
719	20	95.2	220	3	ABG73673	Abg73673 Human col	792	20	95.2	236	2	AAU85591	Aau85591 Lung tumo
720	20	95.2	220	6	ABM65628	Abm65628 Propionib	793	20	95.2	236	2	AAU85591	Aau85591 Lung tumo
721	20	95.2	220	8	ADs24417	Ads24417 Bacterial	794	20	95.2	236	2	ADA83740	Ada83740 Human ASC
722	20	95.2	221	4	ABG60165	Abg60165 Drosophil	795	20	95.2	236	2	ABU69563	Abu69563 Human lun
723	20	95.2	221	4	ABG89423	Abg89423 Plant ful	796	20	95.2	236	2	ABU69563	Abu69563 Human lun
724	20	95.2	222	8	ADx67284	Adx67284 Plant ful	797	20	95.2	236	2	ABU66466	Abu66466 Lung can
725	20	95.2	223	8	ADx68786	Adx68786 Plant ful	798	20	95.2	236	2	ABU66466	Abu66466 Lung can
726	20	95.2	224	5	AAO21843	Aao21843 Pathogeni	799	20	95.2	236	2	ADb11136	Adb11136 Allostoc
727	20	95.2	224	5	ABO09133	Abu09133 Thermotol	800	20	95.2	236	2	ADH47389	Adh47389 Human lun
728	20	95.2	224	5	ABB98517	Abb98517 Pyrococcu	801	20	95.2	236	2	ADH47388	Adh47388 Human lun
729	20	95.2	224	5	ABP53222	Abp53222 Pyrococcu	802	20	95.2	236	2	ABO82278	Abo82278 Pseudomon
730	20	95.2	224	5	AAU78000	Aau78000 Pyrococcu	803	20	95.2	236	2	ADQ16003	Adq16003 Human Mas
731	20	95.2	224	5	ABB97745	Abb97745 P. furios	804	20	95.2	236	2	ADJ21308	Adj21308 Human lun
732	20	95.2	224	6	ABM69759	Abm69759 Phototrab	805	20	95.2	236	2	ADJ21307	Adj21307 Human lun
733	20	95.2	224	7	ADC55436	Adc55436 Protein s	806	20	95.2	236	2	ADJ21307	Adj21307 Human lun
734	20	95.2	224	8	ADR99391	Adr99391 Chlamydia	807	20	95.2	236	2	ABM81462	Abm81462 Tumour-as
735	20	95.2	224	9	AEb40994	Aeb40994 L. pneumo	808	20	95.2	236	2	ADH47388	Adh47388 Human lun
736	20	95.2	224	9	AEb37683	Aeb37683 L. pneumo	809	20	95.2	236	2	ADH47388	Adh47388 Human lun
737	20	95.2	225	3	AAg25205	Aag25205 Arabidops	810	20	95.2	236	2	ADY49491	Ady49491 Human Mas
738	20	95.2	225	9	ABM90636	Abm90636 M. xanthu	811	20	95.2	236	2	ADY27671	Ady27671 Human ach
739	20	95.2	227	7	ADF05977	Adf05977 Bacterial	812	20	95.2	237	5	AEb48888	Aeb48888 N. gonorr
740	20	95.2	227	8	ADU02756	Adu02756 Novel hum	813	20	95.2	237	5	Aau85593	Aau85593 Lung tumo
741	20	95.2	228	3	AAU40130	Aau40130 Arabidops	814	20	95.2	237	6	ABR40861	Abra40861 Zea mays
742	20	95.2	228	3	ABG07467	Abg07467 Arabidops	815	20	95.2	237	6	ABU69565	Abu69565 Human lun
743	20	95.2	228	3	ABG18546	Abg18546 Novel hum	816	20	95.2	237	6	ABU66468	Abu66468 Lung can
744	20	95.2	229	4	ABG18546	Abg18546 Novel hum	817	20	95.2	237	6	ABU66468	Abu66468 Lung can
745	20	95.2	229	4	ABG18546	Abg18546 Novel hum	818	20	95.2	237	6	ADJ21309	Adj21309 Human lun
746	20	95.2	230	3	AAg57483	Aag57483 Arabidops	819	20	95.2	238	4	AAU33967	Aau33967 Staphyloc
747	20	95.2	230	3	ADG60391	Adg60391 Arabidops	820	20	95.2	238	4	ABG15154	Abg15154 Novel hum
748	20	95.2	230	8	ADO17017	Ado17017 Predicted	821	20	95.2	238	4	AAe18734	Aae18734 Human lun
749	20	95.2	230	8	ADS29937	Ads29937 Bacterial	822	20	95.2	238	5	AAU85551	Aau85551 Clone #50
750	20	95.2	231	3	AAW60988	Aaw60988 Streptoco	823	20	95.2	238	5	ABU05365	Abu05365 M. tuberc
751	20	95.2	231	3	AAW13459	Aaw13459 Protein e	824	20	95.2	238	5	ABU05365	Abu05365 M. tuberc
752	20	95.2	231	3	AAW49381	Aaw49381 Arabidops	825	20	95.2	238	6	ABU69523	Abu69523 Human lun
753	20	95.2	231	3	AAW41654	Aaw41654 Arabidops	826	20	95.2	238	6	ABU66426	Abu66426 Lung can
754	20	95.2	231	4	AAW83135	Aaw83135 Protein e	827	20	95.2	238	7	ADB74503	Adb74503 Mycobacte

828	20	95.2	238	7	ADH47332	Adh47332	Humah lun	901	20	95.2	255	3	AAy74711	Aay74711	Neisseria
829	20	95.2	238	8	ADJ21251	Adj21251	Humah lun	902	20	95.2	255	3	AAG21617	Ag21617	Arabidops
830	20	95.2	238	9	ADV70192	Adv70192	Tumor-ass	903	20	95.2	255	3	AAG21660	Ag21660	Arabidops
831	20	95.2	238	9	ADY49493	Ady49493	Human Mas	904	20	95.2	255	3	AAG52164	Ag52164	Arabidops
832	20	95.2	239	2	AAW80690	Aaw80690	S. pneumo	905	20	95.2	255	3	AAG52174	Ag52174	Human pro
833	20	95.2	239	3	AAG07840	Ag07840	Arabidops	906	20	95.2	255	4	AAW93970	Aaw93970	Human sto
834	20	95.2	239	8	ADG93991	Adg93991	Escherich	907	20	95.2	255	6	ABP79581	Abp79581	N. gonorr
835	20	95.2	239	8	ADJ54176	Adj54176	Mara tran	908	20	95.2	255	8	ADT49856	Adt49856	Murine FB
836	20	95.2	239	8	ADO15500	Ado15500	Escherich	909	20	95.2	255	8	ADY22900	Ady22900	Plant ful
837	20	95.2	239	9	AEA32579	Aea32579	Escherich	910	20	95.2	256	4	ABW61922	Abw61922	Drosophil
838	20	95.2	240	4	AAU37860	Aau37860	Streptoco	911	20	95.2	256	8	ADN73617	Adn73617	Thale cre
839	20	95.2	240	4	AAW40842	Aaw40842	Human pol	912	20	95.2	257	2	AAW56493	Aaw56493	Tobacco l
840	20	95.2	240	5	ABP27775	Abp27775	Streptoco	913	20	95.2	257	4	ABG29649	Abg29649	Novel hum
841	20	95.2	240	6	ABU02367	Abu02367	S. pneumo	914	20	95.2	257	6	ABU48422	Abu48422	Novel hum
842	20	95.2	240	6	ABU46228	Abu46228	Protein e	915	20	95.2	257	7	ADH88448	Adh88448	Enterococ
843	20	95.2	240	6	ABU46476	Abu46476	Protein e	916	20	95.2	257	9	ADV09395	Adv09395	TYLCSV ca
844	20	95.2	240	6	ABU44230	Abu44230	Protein e	917	20	95.2	258	4	AAW41766	Aaw41766	Human pol
845	20	95.2	240	8	ADK46455	Adk46455	Streptoco	918	20	95.2	258	6	ABM67981	Abm67981	Phototrab
846	20	95.2	240	8	ADV88518	Adv88518	Streptoco	919	20	95.2	258	8	ADY13339	Ady13339	Plant ful
847	20	95.2	240	8	ADV79771	Adv79771	Streptoco	920	20	95.2	258	8	ABG77446	Abg77446	Selected
848	20	95.2	240	8	ADV81928	Adv81928	Streptoco	921	20	95.2	259	5	ABB54718	Abb54718	Lactococc
849	20	95.2	241	4	AAU38016	Aau38016	Streptoco	922	20	95.2	259	5	ABB54375	Abb54375	Lactococc
850	20	95.2	241	6	ABR52784	AbR52784	Protein s	923	20	95.2	259	5	ABW53792	Abw53792	Lactococc
851	20	95.2	241	7	ADK62144	Adk62144	Diseas e	924	20	95.2	259	5	ABJ11348	Abj11348	yeast sel
852	20	95.2	241	7	ABO64120	AbO64120	Klebsiell	925	20	95.2	259	7	ABM87612	Abm87612	Rice abio
853	20	95.2	241	8	ADQ65226	Adq65226	Novel hum	926	20	95.2	260	4	AAU35725	Aau35725	Helicobac
854	20	95.2	242	3	AAG23135	Aag23135	Arabidops	927	20	95.2	260	4	AAU35725	Aau35725	Helicobac
855	20	95.2	242	4	AAU36571	Aau36571	Staphyloc	928	20	95.2	260	4	AAU35725	Aau35725	Helicobac
856	20	95.2	242	5	ABW54066	Abw54066	Lactococc	929	20	95.2	260	6	ABU30762	Abu30762	Protein e
857	20	95.2	242	6	ABU15970	Abu15970	Protein e	930	20	95.2	260	8	ADQ48612	Adq48612	Helicobac
858	20	95.2	242	6	ABU23340	Abu23340	Protein e	931	20	95.2	260	8	ADS30639	Ads30639	Bacterial
859	20	95.2	242	8	ADK46819	Adk46819	Streptoco	932	20	95.2	261	4	ABW68328	Abw68328	Drosophil
860	20	95.2	242	8	ADK959703	Adk959703	Novel S.	933	20	95.2	261	4	AAW81584	Aaw81584	S. epider
861	20	95.2	242	9	AEA59573	Aea59573	Streptoco	934	20	95.2	261	4	AAW81584	Aaw81584	S. epider
862	20	95.2	244	4	ABG26006	Abg26006	Novel hum	935	20	95.2	261	8	ADL04383	Adl04383	M. catarr
863	20	95.2	244	6	ABR40863	AbR40863	Zea mays	936	20	95.2	262	3	AAW07839	Aaw07839	Arabidops
864	20	95.2	244	8	ADT57497	Adt57497	Plant pol	937	20	95.2	262	7	ADI21247	Adi21247	Novel hum
865	20	95.2	245	8	ABO58572	AbO58572	Humah gen	938	20	95.2	263	6	AAU17415	Aau17415	Novel sig
866	20	95.2	245	8	ADN24505	Adn24505	Bacterial	939	20	95.2	263	6	AAU52589	Aau52589	Human NOV
867	20	95.2	246	3	AAG25204	Ag25204	Arabidops	940	20	95.2	263	7	ADW94123	Adw94123	Human nov
868	20	95.2	246	6	ABR40673	AbR40673	Zea mays	941	20	95.2	263	8	ADS21535	Ads21535	Bacterial
869	20	95.2	246	9	ABW93456	Abw93456	M. xanthu	942	20	95.2	264	5	ABW47511	Abw47511	Listeria
870	20	95.2	247	3	AAG22404	Ag22404	Arabidops	943	20	95.2	264	8	ADU02417	Adu02417	Novel hum
871	20	95.2	247	3	AAG12671	Ag12671	Arabidops	944	20	95.2	266	2	AAU43940	Aau43940	Human pro
872	20	95.2	247	4	ABU53320	Abu53320	Humah cel	945	20	95.2	266	6	ABU45250	Abu45250	Protein e
873	20	95.2	247	4	ABU52991	Abu52991	Humah sig	946	20	95.2	266	6	ABU25437	Abu25437	Protein e
874	20	95.2	247	6	ABJ25658	Abj25658	aspekgill	947	20	95.2	267	8	ADJ50291	Adj50291	Oil-aassoc
875	20	95.2	247	8	ADX89842	Adx89842	Plant ful	948	20	95.2	267	8	ADN27291	Adn27291	Bacterial
876	20	95.2	247	9	ADW18364	Adw18364	Pinu rad	949	20	95.2	267	9	ADW68014	Adw68014	Plant ful
877	20	95.2	248	2	AAU35303	Aau35303	Chlamydia	950	20	95.2	267	9	ABW95402	Abw95402	M. xanthu
878	20	95.2	248	3	AAU75766	Aau75766	Neisseria	951	20	95.2	269	4	ABW63346	Abw63346	Drosophil
879	20	95.2	248	3	AAU75767	Aau75767	Neisseria	952	20	95.2	269	4	AAW40145	Aaw40145	Human pol
880	20	95.2	248	3	AAU75765	Aau75765	Neisseria	953	20	95.2	269	5	AAU80184	Aau80184	Embryonic
881	20	95.2	248	5	ABP27776	Abp27776	Streptoco	954	20	95.2	269	8	ADW09206	Adw09206	Human pro
882	20	95.2	248	6	ABW71601	Abw71601	Staphyloc	955	20	95.2	270	4	AAW82813	Aaw82813	S. epider
883	20	95.2	248	7	ADW26586	Adw26586	Hypether	956	20	95.2	270	6	ABR41385	AbR41385	Human DIT
884	20	95.2	248	8	ADW343965	Adw343965	Bacterial	957	20	95.2	272	2	AAU29201	Aau29201	Amino aci
885	20	95.2	250	8	ADW88498	Adw88498	Plant ful	958	20	95.2	272	3	ABW58955	Abw58955	Breast an
886	20	95.2	250	8	ADW89753	Adw89753	Plant ful	959	20	95.2	272	3	AAW31961	Aaw31961	Arabidops
887	20	95.2	251	3	AAG52165	Ag52165	Arabidops	960	20	95.2	272	5	ABW78032	Abw78032	ITALY, LO
888	20	95.2	251	3	AAG21661	Ag21661	Arabidops	961	20	95.2	272	8	ADW45558	Adw45558	Homologue
889	20	95.2	251	8	ADN99247	Adn99247	Novel hum	962	20	95.2	273	3	AAW17674	Aaw17674	Arabidops
890	20	95.2	251	8	ADW75768	Adw75768	Plant ful	963	20	95.2	273	3	AAW47587	Aaw47587	Arabidops
891	20	95.2	252	3	AAW57482	Aaw57482	Arabidops	964	20	95.2	273	4	AAU58254	Aau58254	Propionib
892	20	95.2	252	3	AAW60390	Aaw60390	Arabidops	965	20	95.2	273	5	ABW93858	Abw93858	Herbicida
893	20	95.2	252	4	AAW79289	Aaw79289	Human pro	966	20	95.2	273	6	ABW54773	Abw54773	Propionib
894	20	95.2	252	7	ABR39143	AbR39143	TRV-RNAI	967	20	95.2	273	6	ABU06044	Abu06044	N. mening
895	20	95.2	252	7	ADM25408	Adm25408	Hypether	968	20	95.2	273	6	ABU37768	Abu37768	Protein e
896	20	95.2	253	3	AAW71553	Aaw71553	Soybean s	969	20	95.2	274	2	AAV17979	Aav17979	Hexulose
897	20	95.2	253	7	ABO66960	AbO66960	Klebsiell	970	20	95.2	275	2	AAV60029	Aav60029	Human end
898	20	95.2	253	8	ADG22675	Adg22675	Cyanophag	971	20	95.2	275	8	ADN24570	Adn24570	Bacterial
899	20	95.2	254	8	ADF58819	Adf58819	Human pol	972	20	95.2	275	8	ADN21810	Adn21810	Bacterial
900	20	95.2	254	8	ADN47979	Adn47979	Thermococ	973	20	95.2	276	3	AAW14069	Aaw14069	Arabidops

974	20	95.2	276	5	ABB91124	Abb91124 Herbicida
975	20	95.2	276	6	ADA34395	Ada34395 Acinetoba
976	20	95.2	276	8	ADX71305	Plant ful
977	20	95.2	277	3	RAG30215	Ag30215 Arabidops
978	20	95.2	277	4	ABG04213	Novel hum
979	20	95.2	278	6	ABM70140	Photorhab
980	20	95.2	278	7	ABO66293	Klebsiell
981	20	95.2	279	2	AA913112	N. gonorr
982	20	95.2	279	6	ABP77390	Abp77390 N. gonorr
983	20	95.2	279	6	ABU50637	Protein e
984	20	95.2	280	3	AA908396	Arabidops
985	20	95.2	280	7	ABO65870	Klebsiell
986	20	95.2	281	7	ABO79112	Pseudomon
987	20	95.2	281	8	ADY09043	Plant ful
988	20	95.2	282	2	AA48330	Human pro
989	20	95.2	282	3	ABG30214	Herbicida
990	20	95.2	282	5	ABB92302	Arabidops
991	20	95.2	282	6	ABP79606	N. gonorr
992	20	95.2	282	7	ADE12782	L. rhamno
993	20	95.2	282	8	ADS28436	Bacterial
994	20	95.2	283	3	AGI14068	Arabidops
995	20	95.2	283	7	ADC86377	Human GPC
996	20	95.2	284	6	ABU44340	Protein e
997	20	95.2	284	8	ADN18215	Bacterial
998	20	95.2	285	5	ABP90597	Chlamydia
999	20	95.2	285	7	ABO67704	Klebsiell
1000	20	95.2	285	8	ADJ48372	Maize oil

ALIGNMENTS

RESULT 1	ADZ71770	standard; peptide; 5 AA.
AC	ADZ71770;	
XX	14-JUL-2005 (first entry)	
DT	p21-derived peptide #355.	
XX	CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;	
KW	neoplasm; cytostatic; pharmaceutical; drug screening.	
XX	Synthetic.	
XX	WO2005040802-A2.	
XX	06-MAY-2005.	
XX	20-OCT-2004; 2004WO-GB004431.	
XX	20-OCT-2003; 2003GB-00024466.	
PR	02-FEB-2004; 2004US-00771242.	
XX	(CYCL-) CYCLACEL LTD.	
PA	Zheleva DI, Fischer PM, McInnes C, Andrews MJI, Chan WC;	
XX	Atkinson GE;	
XX	WPI; 2005-355897/36.	
XX	New peptide inhibitors of cyclin dependent kinases derived from the C-	
PT	terminal region of p21, useful in preparing a medicament for treating a	
PT	proliferative disorder such as cancer.	
XX	Claim 22; Page 103; 112pp; English.	
PS	The invention relates to a peptide or its variant comprising formula: A-	
XX	(B) m -C-(D) n -E , where m or n are each independently 0 or 1; A is a	
CC	natural or unnatural amino acid residue having a side chain comprising at	
CC	least one H-bond acceptor moiety and at least one H-bond donor moiety;	

each of B or D is independently an amino acid residue selected from arginine, glycine, citrulline, glutamine, serine, lysine, asparagine, isoleucine or alanine; C is a natural or unnatural amino acid residue having a branched or unbranched C 1 -C 6 alkylene side chain optionally containing a H-bond donor or a H-bond acceptor moiety; and E is a natural or unnatural amino acid residue having an aryl or heteroaryl side chain. Also described are: a pharmaceutical composition comprising the peptide admixed with a diluent, an excipient or a carrier; an assay for identifying candidate substances capable of binding to a cyclin associated with a G1 control CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact a peptide as defined above, the cyclin, the CDK and the candidate substance, under conditions where, in the absence of the cyclin/CDK interaction, the peptidomimetic would bind to the cyclin; and monitoring any change in the expected binding of the peptide and the cyclin. The assay for identifying compounds that interact a cyclin or a cyclin when complexed with the physiologically relevant CDK comprises: incubating a candidate compound and the peptide and a cyclin or cyclin/CDK complex; and detecting binding of either the candidate compound or the peptide with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay comprises use of a three-dimensional model of a cyclin and a candidate compound. At least one of the assay components is bound to a solid phase. The peptidomimetic is labeled such as to emit a signal when bound to the cyclin. The cyclin is labeled such as to emit a signal when bound to the peptide. One of the assay components is labeled with a fluorescence emitter and the signal is detected using fluorescence polarization techniques. Using a cyclin in a drug screening assay comprises: selecting a candidate compound by performing rational drug design with a three-dimensional model of the cyclin, where the selecting is performed in conjunction with computer modeling; contacting the candidate compound for with the cyclin; and detecting the binding of the candidate compound for the cyclin groove. A potential drug is selected on the basis of its having a greater affinity for the cyclin groove than that of the peptide. The method of detection comprises monitoring G0 and/or G1/S cell cycle, cell cycle-related apoptosis, suppression of E2F transcription factor, hypophosphorylation of cellular pRb, or in vitro anti-proliferative effects. The peptide is useful in preparing a medicament for treating a proliferative disorder, e.g., cancer. The present sequence represents a p21-derived peptide of the invention.	Query Match 95.2%; Score 20; DB 9; Length 5; Best Local Similarity 100.0%; Pred. No. 2e+06; Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 RRLN 4	
Db	1 RRLN 4	
RESULT 2	ADZ72025	
ID	ADZ72025 standard; peptide; 5 AA.	
XX	ADZ72025;	
XX	14-JUL-2005 (first entry)	
XX	p21-derived peptide #610.	
DE	CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;	
KW	neoplasm; cytostatic; pharmaceutical; drug screening.	
XX	Synthetic.	
OS	WO2005040802-A2.	
XX		

PD 06-MAY-2005.  
XX 20-OCT-2004; 2004WO-GB004431.  
XX 20-OCT-2003; 2003GB-00024466.  
XX 02-FEB-2004; 2004US-00771242.  
XX (CYCL-) CYCLACEL LTD.  
XX Zheleva DI, Fischer PM, McInnes C, Andrews MJ, Chan WC;  
XX Atkinson GE;  
XX WPI; 2005-355897/36.  
XX New peptide inhibitors of cyclin dependent kinases derived from the C-  
XX terminal region of p21, useful in preparing a medicament for treating a  
XX proliferative disorder such as cancer.  
XX Example 27; Page 83; 112pp; English.  
XX The invention relates to a peptide or its variant comprising formula: A-  
XX (B) m -C-(D) n -E, where m or n are each independently 0 or 1; A is a  
XX natural or unnatural amino acid residue having a side chain comprising at  
XX least one H-bond acceptor moiety and at least one H-bond donor moiety;  
XX each of B or D is independently an amino acid residue selected from:  
XX arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,  
XX isoleucine or alanine; C is a natural or unnatural amino acid residue  
XX having a branched or unbranched C 1 -C 6 alkylene side chain optionally  
XX containing a H-bond donor or a H-bond acceptor moiety; and E is a natural  
XX or unnatural amino acid residue having an aryl or heteroaryl side chain.  
XX Also described are: a pharmaceutical composition comprising the peptide  
XX admixed with a diluent, an excipient or a carrier; an assay for  
XX identifying candidate substances capable of binding to a cyclin  
XX associated with a G1 control CDK enzyme and/or inhibiting the enzyme;  
XX an assay for identifying compounds that interact a cyclin or a cyclin when  
XX cyclin in a drug screening assay. The assay for identifying candidate  
XX substances capable of binding to a cyclin associated with a G1 control  
XX CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact  
XX a peptide as defined above, the cyclin, the CDK and the candidate  
XX substance, under conditions where, in the absence of the candidate  
XX substance being an inhibitor of interaction of the cyclin/CDK  
XX interaction, the peptidomimetic would bind to the cyclin; and monitoring  
XX any change in the expected binding of the peptide and the cyclin. The  
XX assay for identifying compounds that interact a cyclin or a cyclin when  
XX complexed with the physiologically relevant CDK comprises: incubating a  
XX candidate compound and the peptide and a cyclin or cyclin/CDK complex;  
XX and detecting binding of either the candidate compound or the peptide  
XX with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay  
XX comprises use of a three-dimensional model of a cyclin and a candidate  
XX compound. At least one of the assay components is bound to a solid phase.  
XX The peptidomimetic is labeled such as to emit a signal when bound to the  
XX cyclin. The cyclin is labeled such as to emit a signal when bound to the  
XX peptide. One of the assay components is labeled with a fluorescence  
XX emitter and the signal is detected using fluorescence polarization  
XX techniques. Using a cyclin in a drug screening assay comprises: selecting  
XX a candidate compound by performing rational drug design with a three-  
XX dimensional model of the cyclin, where the selecting is performed in  
XX conjunction with computer modeling; contacting the candidate compound  
XX with the cyclin; and detecting the binding of the candidate compound for  
XX the cyclin groove. A potential drug is selected on the basis of its  
XX having a greater affinity for the cyclin groove than that of the peptide.  
XX The method of detection comprises monitoring G0 and/or G1/S cell cycle,  
XX cell cycle-related apoptosis, suppression of E2F transcription factor,  
XX hypophosphorylation of cellular pRB, or in vitro anti-proliferative  
XX effects. The peptide is useful in preparing a medicament for treating a  
XX proliferative disorder, e.g., cancer. The present sequence represents a  
XX p21-derived peptide of the invention.  
XX Sequence 5 AA;  
XX Query Match 95.2%; Score 20; DB 9; Length 5;  
XX Best Local Similarity 100.0%; Pred. No. 2e+06;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RELN 4  
DB 1 RELN 4  
RESULT 3  
ADZ71763  
ID ADZ71763 standard; peptide; 5 AA.  
XX  
XX AC ADZ71763;  
XX  
XX DT 14-JUL-2005 (first entry)  
XX  
XX DE p21-derived peptide #348.  
XX  
XX KW CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;  
XX neoplasm; cytostatic; pharmaceutical; drug screening.  
XX  
XX OS Synthetic.  
XX  
XX PN WO2005040802-A2.  
XX  
XX PD 06-MAY-2005.  
XX  
XX PF 20-OCT-2004; 2004WO-GB004431.  
XX  
XX PR 20-OCT-2003; 2003GB-00024466.  
XX  
XX PR 02-FEB-2004; 2004US-00771242.  
XX  
XX (CYCL-) CYCLACEL LTD.  
XX  
XX Zheleva DI, Fischer PM, McInnes C, Andrews MJ, Chan WC;  
XX Atkinson GE;  
XX WPI; 2005-355897/36.  
XX  
XX New peptide inhibitors of cyclin dependent kinases derived from the C-  
XX terminal region of p21, useful in preparing a medicament for treating a  
XX proliferative disorder such as cancer.  
XX Claim 22; Page 103; 112pp; English.  
XX The invention relates to a peptide or its variant comprising formula: A-  
XX (B) m -C-(D) n -E, where m or n are each independently 0 or 1; A is a  
XX natural or unnatural amino acid residue having a side chain comprising at  
XX least one H-bond acceptor moiety and at least one H-bond donor moiety;  
XX each of B or D is independently an amino acid residue selected from:  
XX arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,  
XX isoleucine or alanine; C is a natural or unnatural amino acid residue  
XX having a branched or unbranched C 1 -C 6 alkylene side chain optionally  
XX containing a H-bond donor or a H-bond acceptor moiety; and E is a natural  
XX or unnatural amino acid residue having an aryl or heteroaryl side chain  
XX containing a H-bond donor or a H-bond acceptor moiety; and E is a natural  
XX or unnatural amino acid residue having an aryl or heteroaryl side chain.  
XX Also described are: a pharmaceutical composition comprising the peptide  
XX admixed with a diluent, an excipient or a carrier; an assay for  
XX identifying candidate substances capable of binding to a cyclin  
XX associated with a G1 control CDK enzyme and/or inhibiting the enzyme;  
XX an assay for identifying compounds that interact a cyclin or a cyclin when  
XX cyclin in a drug screening assay. The assay for identifying candidate  
XX substances capable of binding to a cyclin associated with a G1 control  
XX CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact  
XX a peptide as defined above, the cyclin, the CDK and the candidate  
XX substance, under conditions where, in the absence of the candidate  
XX substance being an inhibitor of interaction of the cyclin/CDK  
XX interaction, the peptidomimetic would bind to the cyclin; and monitoring  
XX any change in the expected binding of the peptide and the cyclin. The  
XX assay for identifying compounds that interact a cyclin or a cyclin when  
XX complexed with the physiologically relevant CDK comprises: incubating a  
XX candidate compound and the peptide and a cyclin or cyclin/CDK complex;  
XX and detecting binding of either the candidate compound or the peptide  
XX with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay  
XX comprises use of a three-dimensional model of a cyclin and a candidate  
XX compound. At least one of the assay components is bound to a solid phase.  
XX The peptidomimetic is labeled such as to emit a signal when bound to the  
XX cyclin. The cyclin is labeled such as to emit a signal when bound to the  
XX peptide. One of the assay components is labeled with a fluorescence  
XX emitter and the signal is detected using fluorescence polarization  
XX techniques. Using a cyclin in a drug screening assay comprises: selecting  
XX a candidate compound by performing rational drug design with a three-  
XX dimensional model of the cyclin, where the selecting is performed in  
XX conjunction with computer modeling; contacting the candidate compound  
XX with the cyclin; and detecting the binding of the candidate compound for  
XX the cyclin groove. A potential drug is selected on the basis of its  
XX having a greater affinity for the cyclin groove than that of the peptide.  
XX The method of detection comprises monitoring G0 and/or G1/S cell cycle,  
XX cell cycle-related apoptosis, suppression of E2F transcription factor,  
XX hypophosphorylation of cellular pRB, or in vitro anti-proliferative  
XX effects. The peptide is useful in preparing a medicament for treating a  
XX proliferative disorder, e.g., cancer. The present sequence represents a  
XX p21-derived peptide of the invention.  
XX Sequence 5 AA;  
XX Query Match 95.2%; Score 20; DB 9; Length 5;  
XX Best Local Similarity 100.0%; Pred. No. 2e+06;

comprises use of a three-dimensional model of a cyclin and a candidate compound. At least one of the assay components is bound to a solid phase. The peptidomimetic is labeled such as to emit a signal when bound to the cyclin. The cyclin is labeled such as to emit a signal when bound to the peptide. One of the assay components is labeled with a fluorescence emitter and the signal is detected using fluorescence polarization techniques. Using a cyclin in a drug screening assay comprises: selecting a candidate compound by performing rational drug design with a three-dimensional model of the cyclin, where the selecting is performed in conjunction with computer modeling; contacting the candidate compound with the cyclin; and detecting the binding of the candidate compound for having a greater affinity for the cyclin groove than that of the peptide. The method of detection comprises monitoring G0 and/or G1/S cell cycle, cell cycle-related apoptosis, suppression of E2F transcription factor, hypophosphorylation of cellular pRb, or in vitro anti-proliferative effects. The peptide is useful in preparing a medicament for treating a proliferative disorder, e.g., cancer. The present sequence represents a p21-derived peptide of the invention.

Sequence 5 AA;

Query Match 95.2%; Score 20; DB 9; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2e+06; 0; Indels 0; Gaps 0;  
Matches 4; Conservative 0; Mismatches 0;

QY 1 RRLN 4  
|||

Db 1 RRLN 4

RESULT 4

ID ADZ72027 standard; peptide; 5 AA.

AC ADZ72027;

14-JUL-2005 (first entry)

p21-derived peptide #612.

CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;  
neoplasm; cytostatic; pharmaceutical; drug screening.

Synthetic.

WO2005040802-A2.

06-MAY-2005.

20-OCT-2004; 2004WO-GB004431.

20-OCT-2003; 2003GB-00024466.

02-FEB-2004; 2004US-00771242.

(CYCL-) CYCLACEL LTD.

Zheleva DI, Fischer PM, McInnes C, Andrews MJ, Chan WC;  
Atkinson GE;

WPI; 2005-355897/36.

New peptide inhibitors of cyclin dependent kinases derived from the C-terminal region of p21, useful in preparing a medicament for treating a proliferative disorder such as cancer.

Example 27; Page 83; 112pp; English.

The invention relates to a peptide or its variant comprising formula: A-(B) m -C-(D) n -E, where m or n are each independently 0 or 1; A is a natural or unnatural amino acid residue having a side chain comprising at least one H-bond acceptor moiety and at least one H-bond donor moiety; each of B or D is independently an amino acid residue selected from

arginine, glycine, citrulline, glutamine, serine, lysine, asparagine, isoleucine or alanine; C is a natural or unnatural amino acid residue having a branched or unbranched C1-C6 alkylene side chain optionally containing a H-bond donor or a H-bond acceptor moiety; and E is a natural or unnatural amino acid residue having an aryl or heteroaryl side chain. Also described are: a pharmaceutical composition comprising the peptide admixed with a diluent, an excipient or a carrier; an assay for identifying candidate substances capable of binding to a cyclin associated with a G1 control CDK enzyme and/or inhibiting the enzyme; an assay for identifying compounds that interact a cyclin or a cyclin when complexed with the physiologically relevant CDK; and a method of using a cyclin in a drug screening assay. The assay for identifying candidate substances capable of binding to a cyclin associated with a G1 control CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact a peptide as defined above, the cyclin, the CDK and the candidate substance, under conditions where, in the absence of the cyclin/CDK interaction, the peptidomimetic would bind to the cyclin; and monitoring any change in the expected binding of the peptide and the cyclin. The assay for identifying compounds that interact a cyclin or a cyclin when complexed with the physiologically relevant CDK comprises: incubating a candidate compound and the peptide and a cyclin or cyclin/CDK complex with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay comprises use of a three-dimensional model of a cyclin and a candidate compound. At least one of the assay components is bound to a solid phase. The peptidomimetic is labeled such as to emit a signal when bound to the cyclin. The cyclin is labeled such as to emit a signal when bound to the peptide. One of the assay components is labeled with a fluorescence emitter and the signal is detected using fluorescence polarization techniques. Using a cyclin in a drug screening assay comprises: selecting a candidate compound by performing rational drug design with a three-dimensional model of the cyclin, where the selecting is performed in conjunction with computer modeling; contacting the candidate compound for with the cyclin; and detecting the binding of the candidate compound for the cyclin groove. A potential drug is selected on the basis of its having a greater affinity for the cyclin groove than that of the peptide. The method of detection comprises monitoring G0 and/or G1/S cell cycle, cell cycle-related apoptosis, suppression of E2F transcription factor, hypophosphorylation of cellular pRb, or in vitro anti-proliferative effects. The peptide is useful in preparing a medicament for treating a proliferative disorder, e.g., cancer. The present sequence represents a p21-derived peptide of the invention.

Sequence 5 AA;

Query Match 95.2%; Score 20; DB 9; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2e+06; 0; Indels 0; Gaps 0;  
Matches 4; Conservative 0; Mismatches 0;

QY 1 RRLN 4

Db 1 RRLN 4

RESULT 5

ID ADZ72107 standard; peptide; 5 AA.

AC ADZ72107;

14-JUL-2005 (first entry)

p21-derived peptide #692.

CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;  
neoplasm; cytostatic; pharmaceutical; drug screening.

Synthetic.

WO2005040802-A2.

06-MAY-2005.

XX PF 20-OCT-2004; 2004WO-GB004431.  
 XX PR 20-OCT-2003; 2003GB-00024466.  
 XX PR 02-FEB-2004; 2004US-00771242.  
 XX PA (CYCL-) CYCLACEL LTD.  
 XX PI Zheleva DI, Fischer PM, McInnes C, Andrews MJJ, Chan WC;  
 XX PI Atkinson GE;  
 XX DR WPI; 2005-355897/36.  
 XX PT New peptide inhibitors of cyclin dependent kinases derived from the C-  
 XX PT terminal region of p21, useful in preparing a medicament for treating a  
 XX PT proliferative disorder such as cancer.  
 XX PS Example 27; Page 85; 112pp; English.  
 XX CC The invention relates to a peptide or its variant comprising formula: A-  
 CC (B) m -C-(D) n -E, where m or n are each independently 0 or 1; A is a  
 CC natural or unnatural amino acid residue having a side chain comprising at  
 CC least one H-bond acceptor moiety and at least one H-bond donor moiety;  
 CC each of B or D is independently an amino acid residue selected from  
 CC arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,  
 CC isoleucine or alanine; C is a natural or unnatural amino acid residue  
 CC having a branched or unbranched C 1 -C 6 alkylene side chain optionally  
 CC containing a H-bond donor or a H-bond acceptor moiety; and E is a natural  
 CC or unnatural amino acid residue having an aryl or heteroaryl side chain.  
 CC Also described are: a pharmaceutical composition comprising the peptide  
 CC admixed with a diluent, an excipient or a carrier; an assay for  
 CC identifying candidate substances capable of binding to a cyclin  
 CC associated with a G1 control CDK enzyme and/or inhibiting the enzyme; an  
 CC assay for identifying compounds that interact a cyclin or a cyclin when  
 CC complexed with the physiologically relevant CDK; and a method of using a  
 CC cyclin in a drug screening assay. The assay for identifying candidate  
 CC substances capable of binding to a cyclin associated with a G1 control  
 CC CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact  
 CC a peptide as defined above, the cyclin, the CDK and the candidate  
 CC substance, under conditions where, in the absence of the candidate  
 CC substance being an inhibitor of interaction of the cyclin/CDK  
 CC interaction, the peptidomimetic would bind to the cyclin; and monitoring  
 CC any change in the expected binding of the peptide and the cyclin. The  
 CC assay for identifying compounds that interact a cyclin or a cyclin when  
 CC complexed with the physiologically relevant CDK comprises: incubating a  
 CC candidate compound and the peptide and a cyclin or cyclin/CDK complex;  
 CC and detecting binding of either the candidate compound or the peptide  
 CC with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay  
 CC comprises use of a three-dimensional model of a cyclin and a candidate  
 CC compound. At least one of the assay components is bound to a solid phase.  
 CC The peptidomimetic is labeled such as to emit a signal when bound to the  
 CC cyclin. The cyclin is labeled such as to emit a signal when bound to the  
 CC peptide. One of the assay components is labeled with a fluorescence.  
 CC emitter and the signal is detected using fluorescence polarization  
 CC techniques. Using a cyclin in a drug screening assay comprises: selecting  
 CC a candidate compound by performing rational drug design with a three-  
 CC dimensional model of the cyclin, where the selecting is performed in  
 CC conjunction with computer modeling; contacting the candidate compound  
 CC with the cyclin; and detecting the binding of the candidate compound for  
 CC the cyclin groove. A potential drug is selected on the basis of its  
 CC having a greater affinity for the cyclin groove than that of the peptide.  
 CC The method of detection comprises monitoring G0 and/or G1/S cell cycle,  
 CC cell cycle-related apoptosis, suppression of E2F transcription factor,  
 CC hypophosphorylation of cellular pRb, or in vitro anti-proliferative  
 CC effects. The peptide is useful in preparing a medicament for treating a  
 CC proliferative disorder, e.g., cancer. The present sequence represents a  
 CC p21-derived peptide of the invention.  
 XX SQ Sequence 5 AA;

Query Match 95.2%; Score 20; DB 9; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 2e+06; 0; Indels 0; Gaps 0;  
 Matches 4; Conservative 0; Mismatches 0;

OY 1 RRLN 4  
 ID ADZ71617 standard; peptide; 5 AA.  
 XX ADZ71617;  
 AC ADZ71617;  
 XX DT 14-JUL-2005 (First entry)  
 XX DE p21-derived peptide #202.  
 XX KW CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;  
 XX KW neoplasm; cytostatic; pharmaceutical; drug screening.  
 XX OS Synthetic.  
 XX PN WO2005040802-A2.  
 XX PD 06-MAY-2005.  
 XX PF 20-OCT-2004; 2004WO-GB004431.  
 XX PR 20-OCT-2003; 2003GB-00024466.  
 XX PR 02-FEB-2004; 2004US-00771242.  
 XX PA (CYCL-) CYCLACEL LTD.  
 XX PI Zheleva DI, Fischer PM, McInnes C, Andrews MJJ, Chan WC;  
 XX PI Atkinson GE;  
 XX DR WPI; 2005-355897/36.  
 XX PT New peptide inhibitors of cyclin dependent kinases derived from the C-  
 XX PT terminal region of p21, useful in preparing a medicament for treating a  
 XX PT proliferative disorder such as cancer.  
 XX PS Claim 15; Page 98; 112pp; English.  
 XX CC The invention relates to a peptide or its variant comprising formula: A-  
 CC (B) m -C-(D) n -E, where m or n are each independently 0 or 1; A is a  
 CC natural or unnatural amino acid residue having a side chain comprising at  
 CC least one H-bond acceptor moiety and at least one H-bond donor moiety;  
 CC each of B or D is independently an amino acid residue selected from  
 CC arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,  
 CC isoleucine or alanine; C is a natural or unnatural amino acid residue  
 CC having a branched or unbranched C 1 -C 6 alkylene side chain optionally  
 CC containing a H-bond donor or a H-bond acceptor moiety; and E is a natural  
 CC or unnatural amino acid residue having an aryl or heteroaryl side chain.  
 CC Also described are: a pharmaceutical composition comprising the peptide  
 CC admixed with a diluent, an excipient or a carrier; an assay for  
 CC identifying candidate substances capable of binding to a cyclin  
 CC associated with a G1 control CDK enzyme and/or inhibiting the enzyme; an  
 CC assay for identifying compounds that interact a cyclin or a cyclin when  
 CC complexed with the physiologically relevant CDK; and a method of using a  
 CC cyclin in a drug screening assay. The assay for identifying candidate  
 CC substances capable of binding to a cyclin associated with a G1 control  
 CC CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact  
 CC a peptide as defined above, the cyclin, the CDK and the candidate  
 CC substance, under conditions where, in the absence of the candidate  
 CC substance being an inhibitor of interaction of the cyclin/CDK  
 CC interaction, the peptidomimetic would bind to the cyclin; and monitoring  
 CC any change in the expected binding of the peptide and the cyclin. The  
 CC assay for identifying compounds that interact a cyclin or a cyclin when  
 CC complexed with the physiologically relevant CDK comprises: incubating a  
 CC candidate compound and the peptide and a cyclin or cyclin/CDK complex;  
 CC and detecting binding of either the candidate compound or the peptide  
 CC with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay  
 CC comprises use of a three-dimensional model of a cyclin and a candidate  
 CC compound. At least one of the assay components is bound to a solid phase.  
 CC The peptidomimetic is labeled such as to emit a signal when bound to the  
 CC cyclin. The cyclin is labeled such as to emit a signal when bound to the  
 CC peptide. One of the assay components is labeled with a fluorescence.  
 CC emitter and the signal is detected using fluorescence polarization  
 CC techniques. Using a cyclin in a drug screening assay comprises: selecting  
 CC a candidate compound by performing rational drug design with a three-  
 CC dimensional model of the cyclin, where the selecting is performed in  
 CC conjunction with computer modeling; contacting the candidate compound  
 CC with the cyclin; and detecting the binding of the candidate compound for  
 CC the cyclin groove. A potential drug is selected on the basis of its  
 CC having a greater affinity for the cyclin groove than that of the peptide.  
 CC The method of detection comprises monitoring G0 and/or G1/S cell cycle,  
 CC cell cycle-related apoptosis, suppression of E2F transcription factor,  
 CC hypophosphorylation of cellular pRb, or in vitro anti-proliferative  
 CC effects. The peptide is useful in preparing a medicament for treating a  
 CC proliferative disorder, e.g., cancer. The present sequence represents a  
 CC p21-derived peptide of the invention.



compound. At least one of the assay components is bound to a solid phase. The peptidomimetic is labeled such as to emit a signal when bound to the cyclin. The cyclin is labeled such as to emit a signal when bound to the peptide. One of the assay components is labeled with a fluorescence emitter and the signal is detected using fluorescence polarization techniques. Using a cyclin in a drug screening assay comprises: selecting a candidate compound by performing rational drug design with a three-dimensional model of the cyclin, where the selecting is performed in conjunction with computer modeling; contacting the candidate compound with the cyclin; and detecting the binding of the candidate compound for the cyclin groove. A potential drug is selected on the basis of its having a greater affinity for the cyclin groove than that of the peptide. The method of detection comprises monitoring G0 and/or G1/S cell cycle, cell cycle-related apoptosis, suppression of E2F transcription factor, hypophosphorylation of cellular pRb, or in vitro anti-proliferative effects. The peptide is useful in preparing a medicament for treating a proliferative disorder, e.g., cancer. The present sequence represents a p21-derived peptide of the invention.

Query Match 95.2%; Score 20; DB 9; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2e+06; 0; Indels 0; Gaps 0;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
Db 1 RRLN 4

RESULT 7  
ADZ71663  
ID ADZ71663 standard; peptide; 5 AA.  
AC ADZ71663;  
DT 14-JUL-2005 (first entry)  
XX p21-derived peptide #248.  
DE CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;  
KW neoplasm; cytostatic; pharmaceutical; drug screening.  
XX Synthetic.  
OS  
XX  
FN W02005040802-A2.  
XX  
PD 06-MAY-2005.  
XX  
XX 20-OCT-2004; 2004WO-GB004431.  
XX  
XX 20-OCT-2003; 2003GB-00024466.  
PR 02-FEB-2004; 2004US-00771242.  
XX  
XX (CYCL-) CYCLACEL LTD.  
XX  
XX Zheleva DI, Fischer PM, McInnes C, Andrews MJ, Chan WC;  
PI Atkinson GE;  
XX  
XX WPI; 2005-355897/36.  
XX  
XX New peptide inhibitors of cyclin dependent kinases derived from the C-terminal region of p21, useful in preparing a medicament for treating a proliferative disorder such as cancer.  
XX  
XX Claim 20; Page 100; 112pp; English.  
XX  
XX The invention relates to a peptide or its variant comprising formula: A-(B) m .C-(D) n -E, where m or n are each independently 0 or 1; A is a natural or unnatural amino acid residue having a side chain comprising at least one H-bond acceptor moiety and at least one H-bond donor moiety; each of B or D is independently an amino acid residue selected from arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,

isoleucine or alanine; C is a natural or unnatural amino acid residue having a branched or unbranched C 1 -C 6 alkylene side chain optionally containing a H-bond donor or a H-bond acceptor moiety; and E is a natural or unnatural amino acid residue having an aryl or heteroaryl side chain. Also described are: a pharmaceutical composition comprising the peptide admixed with a diluent, an excipient or a carrier; an assay for identifying candidate substances capable of binding to a cyclin associated with a G1 control CDK enzyme and/or inhibiting the enzyme; an assay for identifying compounds that interact a cyclin or a cyclin when complexed with the physiologically relevant CDK; and a method of using a cyclin in a drug screening assay. The assay for identifying candidate substances capable of binding to a cyclin associated with a G1 control CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact a peptide as defined above, the cyclin, the CDK and the candidate substance, under conditions where, in the absence of the candidate substance, under conditions where, in the absence of the cyclin/CDK interaction, the peptidomimetic would bind to the cyclin; and monitoring any change in the expected binding of the peptide and the cyclin. The assay for identifying compounds that interact a cyclin or a cyclin when complexed with the physiologically relevant CDK comprises: incubating a candidate compound and the peptide and a cyclin or cyclin/CDK complex; and detecting binding of either the candidate compound or the peptide with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay comprises use of a three-dimensional model of a cyclin and a candidate compound. At least one of the assay components is bound to a solid phase. The peptidomimetic is labeled such as to emit a signal when bound to the cyclin. The cyclin is labeled such as to emit a signal when bound to the peptide. One of the assay components is labeled with a fluorescence emitter and the signal is detected using fluorescence polarization techniques. Using a cyclin in a drug screening assay comprises: selecting a candidate compound by performing rational drug design with a three-dimensional model of the cyclin, where the selecting is performed in conjunction with computer modeling; contacting the candidate compound for the cyclin groove. A potential drug is selected on the basis of its having a greater affinity for the cyclin groove than that of the peptide. The method of detection comprises monitoring G0 and/or G1/S cell cycle, cell cycle-related apoptosis, suppression of E2F transcription factor, hypophosphorylation of cellular pRb, or in vitro anti-proliferative effects. The peptide is useful in preparing a medicament for treating a proliferative disorder, e.g., cancer. The present sequence represents a p21-derived peptide of the invention.

Sequence 5 AA;

Query Match 95.2%; Score 20; DB 9; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
Db 1 RRLN 4

RESULT 8  
ADZ71764  
ID ADZ71764 standard; peptide; 5 AA.  
XX  
AC ADZ71764;  
XX  
DT 14-JUL-2005 (first entry)  
XX  
DE p21-derived peptide #349.  
XX  
KW CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;  
KW neoplasm; cytostatic; pharmaceutical; drug screening.  
XX  
OS Synthetic.  
XX  
XX W02005040802-A2.  
XX  
PD 06-MAY-2005.  
XX





Fri Jan 27 10:08:32 2006

CC The peptidomimetic is labeled such as to emit a signal when bound to the  
 CC cyclin. The cyclin is labeled such as to emit a signal when bound to the  
 CC peptide. One of the assay components is labeled with a fluorescence  
 CC emitter and the signal is detected using fluorescence polarization  
 CC techniques. Using a cyclin in a drug screening assay comprises: selecting  
 CC a candidate compound by performing rational drug design with a three-  
 CC dimensional model of the cyclin, where the selecting is performed in  
 CC conjunction with computer modeling; contacting the candidate compound for  
 CC with the cyclin; and detecting the binding of the candidate compound for  
 CC the cyclin groove. A potential drug is selected on the basis of its  
 CC having a greater affinity for the cyclin groove than that of the peptide.  
 CC The method of detection comprises monitoring G0 and/or G1/S cell cycle,  
 CC cell cycle-related apoptosis, suppression of E2F transcription factor,  
 CC hypophosphorylation of cellular pRb, or in vitro anti-proliferative  
 CC effects. The peptide is useful in preparing a medicament for treating a  
 CC proliferative disorder, e.g., cancer. The present sequence represents a  
 CC p21-derived peptide of the invention.

XX Sequence 5 AA;  
 SQ Query Match 95.2%; Score 20; DB 9; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 2e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
 DB 1 RRLN 4  
 RESULT 10  
 ID ADZ71746  
 AC ADZ71746;  
 DT 14-JUL-2005 (first entry)  
 DE p21-derived peptide #331.  
 DE CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;  
 KW neoplasm; cytostatic; pharmaceutical; drug screening.  
 XX Synthetic.  
 XX WO2005040802-A2.  
 XX 06-MAY-2005.  
 XX 20-OCT-2004; 2004WO-GB004431.  
 XX 20-OCT-2003; 2003GB-00024466.  
 PR 02-FEB-2004; 2004US-00771242.  
 XX (CYCL-) CYCLACEL LTD.  
 XX Zheleva DI, Fischer PM, McInnes C, Andrews MJ, Chan WC;  
 PI Atkinson GE;  
 PI WPI; 2005-355897/36.  
 DR New peptide inhibitors of cyclin dependent kinases derived from the C-  
 XX terminal region of p21, useful in preparing a medicament for treating a  
 XX proliferative disorder such as cancer.  
 XX Claim 21; Page 102; 112pp; English.

XX The invention relates to a peptide or its variant comprising formula: A-  
 CC (B) m -C-(D) n -E, where m or n are each independently 0 or 1; A is a  
 CC natural or unnatural amino acid residue having a side chain comprising at  
 CC least one H-bond acceptor moiety and at least one H-bond donor moiety;  
 CC each of B or D is independently an amino acid residue selected from  
 CC arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,  
 CC isoleucine or alanine; C is a natural or unnatural amino acid residue

CC having a branched or unbranched C 1 -C 6 alkylene side chain optionally  
 CC containing a H-bond donor or a H-bond acceptor moiety; and E is a natural  
 CC or unnatural amino acid residue having an aryl or heteroaryl side chain.  
 CC Also described are: a pharmaceutical composition comprising the peptide  
 CC admixed with a diluent, an excipient or a carrier; an assay for  
 CC identifying candidate substances capable of binding to a cyclin  
 CC associated with a G1 control CDK enzyme and/or inhibiting the enzyme; an  
 CC assay for identifying compounds that interact a cyclin or a cyclin when  
 CC complexed with the physiologically relevant CDK; and a method of using a  
 CC cyclin in a drug screening assay. The assay for identifying candidate  
 CC substances capable of binding to a cyclin associated with a G1 control  
 CC CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact  
 CC a peptide as defined above, the cyclin, the CDK and the candidate  
 CC substance, under conditions where, in the absence of the candidate  
 CC substance, the peptidomimetic would bind to the cyclin; and monitoring  
 CC interaction, the peptidomimetic binding of the peptide and the cyclin when  
 CC any change in the expected binding of the peptide and the cyclin is  
 CC observed. At least one of the assay components is bound to a solid phase.  
 CC The peptidomimetic is labeled such as to emit a signal when bound to the  
 CC cyclin. The cyclin is labeled such as to emit a signal when bound to the  
 CC peptide. One of the assay components is labeled with a fluorescence  
 CC emitter and the signal is detected using fluorescence polarization  
 CC techniques. Using a cyclin in a drug screening assay comprises: selecting  
 CC a candidate compound by performing rational drug design with a three-  
 CC dimensional model of the cyclin, where the selecting is performed in  
 CC conjunction with computer modeling; contacting the candidate compound for  
 CC with the cyclin; and detecting the binding of the candidate compound for  
 CC the cyclin groove. A potential drug is selected on the basis of its  
 CC having a greater affinity for the cyclin groove than that of the peptide.  
 CC The method of detection comprises monitoring G0 and/or G1/S cell cycle,  
 CC cell cycle-related apoptosis, suppression of E2F transcription factor,  
 CC hypophosphorylation of cellular pRb, or in vitro anti-proliferative  
 CC effects. The peptide is useful in preparing a medicament for treating a  
 CC proliferative disorder, e.g., cancer. The present sequence represents a  
 CC p21-derived peptide of the invention.

XX Sequence 5 AA;  
 SQ Query Match 95.2%; Score 20; DB 9; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 2e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
 DB 1 RRLN 4  
 RESULT 11  
 ID ADZ71618  
 AC ADZ71618;  
 DT 14-JUL-2005 (first entry)  
 DE p21-derived peptide #203.  
 DE CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;  
 KW neoplasm; cytostatic; pharmaceutical; drug screening.  
 XX Synthetic.  
 XX WO2005040802-A2.  
 XX 06-MAY-2005.  
 XX 20-OCT-2004; 2004WO-GB004431.

XX 20-OCT-2003; 2003GB-00024466.  
PR 02-FEB-2004; 2004US-00771242.  
PA (CYCL-) CYCLACEL LTD.  
XX Zheleva DI, Fischer PM, McInnes C, Andrews MJ, Chan WC;  
PI Atkinson GE;  
XX WPI; 2005-355897/36.  
XX New peptide inhibitors of cyclin dependent kinases derived from the C-  
PT terminal region of p21, useful in preparing a medicament for treating a  
PT proliferative disorder such as cancer.  
XX Claim 15; Page 98; 112pp; English.  
XX The invention relates to a peptide or its variant comprising formula: A-  
CC (B) m -C-(D) n -E, where m or n are each independently 0 or 1; A is a  
CC natural or unnatural amino acid residue having a side chain comprising at  
CC least one H-bond acceptor moiety and at least one H-bond donor moiety;  
CC each of B or D is independently an amino acid residue selected from:  
CC arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,  
CC isoleucine or alanine; C is a natural or unnatural amino acid residue  
CC containing a H-bond donor or a H-bond acceptor moiety; and E is a natural  
CC or unnatural amino acid residue having an aryl or heteroaryl side chain.  
CC Also described are: a pharmaceutical composition comprising the peptide  
CC admixed with a diluent, an excipient or a carrier; an assay for  
CC identifying candidate substances capable of binding to a cyclin  
CC associated with a G1 control CDK enzyme and/or inhibiting the enzyme; an  
CC assay for identifying compounds that interact a cyclin or a cyclin when  
CC complexed with the physiologically relevant CDK; and a method of using a  
CC cyclin in a drug screening assay. The assay for identifying candidate  
CC substances capable of binding to a cyclin associated with a G1 control  
CC CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact  
CC a peptide as defined above, the cyclin, the CDK and the candidate  
CC substance, under conditions where, in the absence of the candidate  
CC substance being an inhibitor of interaction of the cyclin/CDK  
CC interaction, the peptidomimetic would bind to the cyclin; and monitoring  
CC any change in the expected binding of the peptide and the cyclin. The  
CC assay for identifying compounds that interact a cyclin or a cyclin when  
CC complexed with the physiologically relevant CDK comprises: incubating a  
CC candidate compound and the peptide and a cyclin or cyclin/CDK complex;  
CC and detecting binding of either the candidate compound or the peptide  
CC with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay  
CC comprises use of a three-dimensional model of a cyclin and a candidate  
CC compound. At least one of the assay components is bound to a solid phase.  
CC The peptidomimetic is labeled such as to emit a signal when bound to the  
CC cyclin. The cyclin is labeled such as to emit a signal when bound to the  
CC peptide. One of the assay components is labeled with a fluorescence  
CC emitter and the signal is detected using fluorescence polarization  
CC techniques. Using a cyclin in a drug screening assay comprises: selecting  
CC a candidate compound by performing rational drug design with a three-  
CC dimensional model of the cyclin, where the selecting is performed in  
CC conjunction with computer modeling; contacting the candidate compound  
CC with the cyclin; and detecting the binding of the candidate compound for  
CC the cyclin groove. A potential drug is selected on the basis of its  
CC having a greater affinity for the cyclin groove than that of the peptide.  
CC The method of detection comprises monitoring G0 and/or G1/S cell cycle,  
CC cell cycle-related apoptosis, suppression of E2F transcription factor,  
CC hypophosphorylation of cellular pRb, or in vitro anti-proliferative  
CC effects. The peptide is useful in preparing a medicament for treating a  
CC proliferative disorder, e.g., cancer. The present sequence represents a  
CC p21-derived peptide of the invention.

XX Sequence 5 AA;  
SQ Query Match 95.2%; Score 20; DB 9; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RRLN 4

DB RESULT 12  
AD271662  
ID ADZ71662 standard; peptide; 5 AA.  
XX AC ADZ71662;  
XX DT 14-JUL-2005 (first entry)  
XX DE p21-derived peptide #247.  
XX KW CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;  
KW neoplasm; cytostatic; pharmaceutical; drug screening.  
XX OS Synthetic.  
XX PN WO2005040802-A2.  
XX PD 06-MAY-2005.  
XX PF 20-OCT-2004; 2004WO-GB004431.  
XX PR 20-OCT-2003; 2003GB-00024466.  
XX PR 02-FEB-2004; 2004US-00771242.  
XX (CYCL-) CYCLACEL LTD.  
XX Zheleva DI, Fischer PM, McInnes C, Andrews MJ, Chan WC;  
PI Atkinson GE;  
XX WPI; 2005-355897/36.  
XX New peptide inhibitors of cyclin dependent kinases derived from the C-  
PT terminal region of p21, useful in preparing a medicament for treating a  
PT proliferative disorder such as cancer.  
PT Claim 20; Page 100; 112pp; English.  
XX The invention relates to a peptide or its variant comprising formula: A-  
CC (B) m -C-(D) n -E, where m or n are each independently 0 or 1; A is a  
CC natural or unnatural amino acid residue having a side chain comprising at  
CC least one H-bond acceptor moiety and at least one H-bond donor moiety;  
CC each of B or D is independently an amino acid residue selected from:  
CC arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,  
CC isoleucine or alanine; C is a natural or unnatural amino acid residue  
CC containing a H-bond donor or a H-bond acceptor moiety; and E is a natural  
CC or unnatural amino acid residue having an aryl or heteroaryl side chain.  
CC Also described are: a pharmaceutical composition comprising the peptide  
CC admixed with a diluent, an excipient or a carrier; an assay for  
CC identifying candidate substances capable of binding to a cyclin  
CC associated with a G1 control CDK enzyme and/or inhibiting the enzyme; an  
CC assay for identifying compounds that interact a cyclin or a cyclin when  
CC complexed with the physiologically relevant CDK; and a method of using a  
CC cyclin in a drug screening assay. The assay for identifying candidate  
CC substances capable of binding to a cyclin associated with a G1 control  
CC CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact  
CC a peptide as defined above, the cyclin, the CDK and the candidate  
CC substance, under conditions where, in the absence of the candidate  
CC substance being an inhibitor of interaction of the cyclin/CDK  
CC interaction, the peptidomimetic would bind to the cyclin; and monitoring  
CC any change in the expected binding of the peptide and the cyclin. The  
CC assay for identifying compounds that interact a cyclin or a cyclin when  
CC complexed with the physiologically relevant CDK comprises: incubating a  
CC candidate compound and the peptide and a cyclin or cyclin/CDK complex;  
CC and detecting binding of either the candidate compound or the peptide  
CC with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay  
CC comprises use of a three-dimensional model of a cyclin and a candidate  
CC compound. At least one of the assay components is bound to a solid phase.  
CC The peptidomimetic is labeled such as to emit a signal when bound to the  
CC cyclin. The cyclin is labeled such as to emit a signal when bound to the  
CC peptide. One of the assay components is labeled with a fluorescence  
CC emitter and the signal is detected using fluorescence polarization  
CC techniques. Using a cyclin in a drug screening assay comprises: selecting  
CC a candidate compound by performing rational drug design with a three-  
CC dimensional model of the cyclin, where the selecting is performed in  
CC conjunction with computer modeling; contacting the candidate compound  
CC with the cyclin; and detecting the binding of the candidate compound for  
CC the cyclin groove. A potential drug is selected on the basis of its  
CC having a greater affinity for the cyclin groove than that of the peptide.  
CC The method of detection comprises monitoring G0 and/or G1/S cell cycle,  
CC cell cycle-related apoptosis, suppression of E2F transcription factor,  
CC hypophosphorylation of cellular pRb, or in vitro anti-proliferative  
CC effects. The peptide is useful in preparing a medicament for treating a  
CC proliferative disorder, e.g., cancer. The present sequence represents a  
CC p21-derived peptide of the invention.

CC cyclin. The cyclin is labeled such as to emit a signal when bound to the  
 CC peptide. One of the assay components is labeled with a fluorescence  
 CC emitter and the signal is detected using fluorescence polarization  
 CC techniques. Using a cyclin in a drug screening assay comprises: selecting  
 CC a candidate compound by performing rational drug design with a three-  
 CC dimensional model of the cyclin, where the selecting is performed in  
 CC conjunction with computer modeling; contacting the candidate compound for  
 CC with the cyclin; and detecting the binding of the candidate compound for  
 CC the cyclin groove. A potential drug is selected on the basis of its  
 CC having a greater affinity for the cyclin groove than that of the peptide.  
 CC The method of detection comprises monitoring G0 and/or G1/S cell cycle,  
 CC cell cycle-related apoptosis, suppression of E2F transcription factor,  
 CC hypophosphorylation of cellular pRb, or in vitro anti-proliferative  
 CC effects. The peptide is useful in preparing a medicament for treating a  
 CC proliferative disorder, e.g., cancer. The present sequence represents a  
 CC p21-derived peptide of the invention.

XX Sequence 5 AA;

Query Match 95.2%; Score 20; DB 9; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 2e+06; 0; Indels 0; Gaps 0;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
 Db ||||  
 1 RRLN 4

RESULT 13

ADZ711664  
 ID ADZ711664 standard; peptide; 5 AA.

AC ADZ711664;

XX 14-JUL-2005 (first entry)

XX p21-derived peptide #249.

XX CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;  
 KW neoplasm; cytostatic; pharmaceutical; drug screening.

XX Synthetic.

XX WO2005040802-A2.

XX 06-MAY-2005.

XX 20-OCT-2004; 2004WO-GB004431.

XX 20-OCT-2003; 2003GB-00024466.

PR 02-FEB-2004; 2004US-00771242.

XX (CYCL-) CYCLACEL LTD.

XX Zheleva DI, Fischer PM, McInnes C, Andrews MJ, Chan WC;

PI Atkinson GE;

XX WPI; 2005-355897/36.

XX New peptide inhibitors of cyclin dependent kinases derived from the C-  
 PT terminal region of p21, useful in preparing a medicament for treating a  
 PT proliferative disorder such as cancer.

XX Claim 20; Page 100; 112pp; English.

XX The invention relates to a peptide or its variant comprising formula: A-  
 CC (B) m -C-(D) n -E, where m or n are each independently 0 or 1; A is a  
 CC natural or unnatural amino acid residue having a side chain comprising at  
 CC least one H-bond acceptor moiety and at least one H-bond donor moiety;  
 CC each of B or D is independently an amino acid residue selected from  
 CC arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,  
 CC isoleucine or alanine; C is a natural or unnatural amino acid residue  
 CC having a branched or unbranched C-1-C-6 alkylene side chain optionally

CC containing a H-bond donor or a H-bond acceptor moiety; and E is a natural  
 CC or unnatural amino acid residue having an aryl or heteroaryl side chain.  
 CC Also described are: a pharmaceutical composition comprising the peptide  
 CC admixed with a diluent, an excipient or a carrier; an assay for  
 CC identifying candidate substances capable of binding to a cyclin  
 CC associated with a G1 control CDK enzyme and/or inhibiting the enzyme; an  
 CC assay for identifying compounds that interact a cyclin or a cyclin when  
 CC complexed with the physiologically relevant CDK; and a method of using a  
 CC cyclin in a drug screening assay. The assay for identifying candidate  
 CC substances capable of binding to a cyclin associated with a G1 control  
 CC CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact  
 CC a peptide as defined above, the cyclin, the CDK and the candidate  
 CC substance, under conditions where, in the absence of the candidate  
 CC substance being an inhibitor of interaction of the cyclin/CDK  
 CC interaction, the peptidomimetic would bind to the cyclin; and monitoring  
 CC any change in the expected binding of the peptide and the cyclin. The  
 CC assay for identifying compounds that interact a cyclin or a cyclin when  
 CC complexed with the physiologically relevant CDK comprises: incubating a  
 CC candidate compound and the peptide and a cyclin or cyclin/CDK complex;  
 CC and detecting binding of either the candidate compound or the peptide  
 CC with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay  
 CC comprises use of a three-dimensional model of a cyclin and a candidate  
 CC compound. At least one of the assay components is bound to a solid phase.  
 CC The peptidomimetic is labeled such as to emit a signal when bound to the  
 CC cyclin. The cyclin is labeled such as to emit a signal when bound to the  
 CC peptide. One of the assay components is labeled with a fluorescence  
 CC emitter and the signal is detected using fluorescence polarization  
 CC techniques. Using a cyclin in a drug screening assay comprises: selecting  
 CC a candidate compound by performing rational drug design with a three-  
 CC dimensional model of the cyclin, where the selecting is performed in  
 CC conjunction with computer modeling; contacting the candidate compound for  
 CC with the cyclin; and detecting the binding of the candidate compound for  
 CC the cyclin groove. A potential drug is selected on the basis of its  
 CC having a greater affinity for the cyclin groove than that of the peptide.  
 CC The method of detection comprises monitoring G0 and/or G1/S cell cycle,  
 CC cell cycle-related apoptosis, suppression of E2F transcription factor,  
 CC hypophosphorylation of cellular pRb, or in vitro anti-proliferative  
 CC effects. The peptide is useful in preparing a medicament for treating a  
 CC proliferative disorder, e.g., cancer. The present sequence represents a  
 CC p21-derived peptide of the invention.

XX Sequence 5 AA;

Query Match 95.2%; Score 20; DB 9; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 2e+06; 0; Indels 0; Gaps 0;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
 Db ||||  
 1 RRLN 4

RESULT 14

ADZ72026  
 ID ADZ72026 standard; peptide; 5 AA.

XX AC ADZ72026;

XX 14-JUL-2005 (first entry)

XX p21-derived peptide #611.

XX CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;  
 KW neoplasm; cytostatic; pharmaceutical; drug screening.

XX Synthetic.

XX WO2005040802-A2.

XX 06-MAY-2005.

XX 20-OCT-2004; 2004WO-GB004431.

PR 20-OCT-2003; 2003GB-00024466.  
PR 02-FEB-2004; 2004US-00771242.  
XX (CYCL-) CYCLACEL LTD.  
XX Zheleva DI, Fischer PM, McInnes C, Andrews MJJ, Chan WC;  
XX Atkinson GB;  
XX WPI; 2005-355897/36.  
XX  
XX New peptide inhibitors of cyclin dependent kinases derived from the C-  
XX terminal region of p21, useful in preparing a medicament for treating a  
XX proliferative disorder such as cancer.  
XX  
XX Example 27; Page 83; 112pp; English.  
XX  
XX The invention relates to a peptide or its variant comprising formula: A-  
XX (B) m -C-(D) n -E, where m or n are each independently 0 or 1; A is a  
XX natural or unnatural amino acid residue having a side chain comprising at  
XX least one H-bond acceptor moiety and at least one H-bond donor moiety;  
XX each of B or D is independently an amino acid residue selected from  
XX arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,  
XX isoleucine or alanine; C is a natural or unnatural amino acid residue  
XX having a branched or unbranched C1-C6 alkylene side chain optionally  
XX containing a H-bond donor or a H-bond acceptor moiety; and E is a natural  
XX or unnatural amino acid residue having an aryl or heteroaryl side chain.  
XX Also described are: a pharmaceutical composition comprising the peptide  
XX admixed with a diluent, an excipient or a carrier; an assay for  
XX identifying candidate substances capable of binding to a cyclin  
XX associated with a G1 control CDK enzyme and/or inhibiting the enzyme; an  
XX assay for identifying compounds that interact a cyclin or a cyclin when  
XX cyclin in a drug screening assay. The assay for identifying candidate  
XX substances capable of binding to a cyclin associated with a G1 control  
XX CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact  
XX a peptide as defined above, the cyclin, the CDK and the candidate  
XX substance, under conditions where, in the absence of the candidate  
XX substance being an inhibitor of interaction of the cyclin/CDK  
XX interaction, the peptidomimetic would bind to the cyclin; and monitoring  
XX any change in the expected binding of the peptide and the cyclin. The  
XX assay for identifying compounds that interact a cyclin or a cyclin when  
XX cyclin in a drug screening assay comprises: incubating a  
XX candidate compound and the peptide and a cyclin or cyclin/CDK complex;  
XX and detecting binding of either the candidate compound or the peptide  
XX with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay  
XX comprises use of a three-dimensional model of a cyclin and a candidate  
XX compound. At least one of the assay components is bound to a solid phase.  
XX The peptidomimetic is labeled such as to emit a signal when bound to the  
XX cyclin. The cyclin is labeled such as to emit a signal when bound to the  
XX peptide. One of the assay components is labeled with a fluorescence  
XX emitter and the signal is detected using fluorescence polarization  
XX techniques. Using a cyclin in a drug screening assay comprises: selecting  
XX a candidate compound by performing rational drug design with a three-  
XX dimensional model of the cyclin, where the selecting is performed in  
XX conjunction with computer modeling; contacting the candidate compound  
XX with the cyclin; and detecting the binding of the candidate compound for  
XX the cyclin groove. A potential drug is selected on the basis of its  
XX having a greater affinity for the cyclin groove than that of the peptide.  
XX The method of detection comprises monitoring G0 and/or G1/S cell cycle,  
XX cell cycle-related apoptosis, suppression of E2F transcription factor,  
XX hypophosphorylation of cellular pRb, or in vitro anti-proliferative  
XX effects. The peptide is useful in preparing a medicament for treating a  
XX proliferative disorder, e.g., cancer. The present sequence represents a  
XX p21-derived peptide of the invention.  
XX  
XX Sequence 5 AA;  
XX  
XX Query Match 95.2%; Score 20; DB 9; Length 5;  
XX Best Local Similarity 100.0%; Pred. No. 2e+06;  
XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX 1 RRLN 4  
XX ||||

Db 1 RRLN 4  
RESULT 15  
ADZ72106  
ID ADZ72106 standard; peptide; 5 AA.  
XX  
XX AC ADZ72106;  
XX  
XX DT 14-JUL-2005 (first entry)  
XX  
XX DE p21-derived peptide #691.  
XX  
XX KW CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;  
XX neoplasm; cytostatic; pharmaceutical; drug screening.  
XX  
XX OS Synthetic.  
XX  
XX PN WO2005040802-A2.  
XX  
XX PD 06-MAY-2005.  
XX  
XX PF 20-OCT-2004; 2004WO-GB004431.  
XX  
XX PR 20-OCT-2003; 2003GB-00024466.  
XX  
XX PR 02-FEB-2004; 2004US-00771242.  
XX (CYCL-) CYCLACEL LTD.  
XX  
XX Zheleva DI, Fischer PM, McInnes C, Andrews MJJ, Chan WC;  
XX Atkinson GB;  
XX WPI; 2005-355897/36.  
XX  
XX New peptide inhibitors of cyclin dependent kinases derived from the C-  
XX terminal region of p21, useful in preparing a medicament for treating a  
XX proliferative disorder such as cancer.  
XX  
XX Example 27; Page 85; 112pp; English.  
XX  
XX The invention relates to a peptide or its variant comprising formula: A-  
XX (B) m -C-(D) n -E, where m or n are each independently 0 or 1; A is a  
XX natural or unnatural amino acid residue having a side chain comprising at  
XX least one H-bond acceptor moiety and at least one H-bond donor moiety;  
XX each of B or D is independently an amino acid residue selected from  
XX arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,  
XX isoleucine or alanine; C is a natural or unnatural amino acid residue  
XX having a branched or unbranched C1-C6 alkylene side chain optionally  
XX containing a H-bond donor or a H-bond acceptor moiety; and E is a natural  
XX or unnatural amino acid residue having an aryl or heteroaryl side chain.  
XX Also described are: a pharmaceutical composition comprising the peptide  
XX admixed with a diluent, an excipient or a carrier; an assay for  
XX identifying candidate substances capable of binding to a cyclin  
XX associated with a G1 control CDK enzyme and/or inhibiting the enzyme; an  
XX assay for identifying compounds that interact a cyclin or a cyclin when  
XX cyclin in a drug screening assay. The assay for identifying candidate  
XX substances capable of binding to a cyclin associated with a G1 control  
XX CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact  
XX a peptide as defined above, the cyclin, the CDK and the candidate  
XX substance, under conditions where, in the absence of the candidate  
XX substance being an inhibitor of interaction of the cyclin/CDK  
XX interaction, the peptidomimetic would bind to the cyclin; and monitoring  
XX any change in the expected binding of the peptide and the cyclin. The  
XX assay for identifying compounds that interact a cyclin or a cyclin when  
XX cyclin in a drug screening assay comprises: incubating a  
XX candidate compound and the peptide and a cyclin or cyclin/CDK complex;  
XX and detecting binding of either the candidate compound or the peptide  
XX with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay  
XX comprises use of a three-dimensional model of a cyclin and a candidate  
XX compound. At least one of the assay components is bound to a solid phase.  
XX The peptidomimetic is labeled such as to emit a signal when bound to the  
XX cyclin. The cyclin is labeled such as to emit a signal when bound to the  
XX peptide. One of the assay components is labeled with a fluorescence  
XX emitter and the signal is detected using fluorescence polarization  
XX techniques. Using a cyclin in a drug screening assay comprises: selecting  
XX a candidate compound by performing rational drug design with a three-  
XX dimensional model of the cyclin, where the selecting is performed in  
XX conjunction with computer modeling; contacting the candidate compound  
XX with the cyclin; and detecting the binding of the candidate compound for  
XX the cyclin groove. A potential drug is selected on the basis of its  
XX having a greater affinity for the cyclin groove than that of the peptide.  
XX The method of detection comprises monitoring G0 and/or G1/S cell cycle,  
XX cell cycle-related apoptosis, suppression of E2F transcription factor,  
XX hypophosphorylation of cellular pRb, or in vitro anti-proliferative  
XX effects. The peptide is useful in preparing a medicament for treating a  
XX proliferative disorder, e.g., cancer. The present sequence represents a  
XX p21-derived peptide of the invention.  
XX  
XX Sequence 5 AA;  
XX  
XX Query Match 95.2%; Score 20; DB 9; Length 5;  
XX Best Local Similarity 100.0%; Pred. No. 2e+06;  
XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX 1 RRLN 4  
XX ||||

CC peptide. One of the assay components is labeled with a fluorescence  
 CC emitter and the signal is detected using fluorescence polarization  
 CC techniques. Using a cyclin in a drug screening assay comprises: selecting  
 CC a candidate compound by performing rational drug design with a three-  
 CC dimensional model of the cyclin, where the selecting is performed in  
 CC conjunction with computer modeling; contacting the candidate compound for  
 CC with the cyclin; and detecting the binding of the candidate compound for  
 CC the cyclin groove. A potential drug is selected on the basis of its  
 CC having a greater affinity for the cyclin groove than that of the peptide.  
 CC The method of detection comprises monitoring G0 and/or G1/S cell cycle,  
 CC cell cycle-related apoptosis, suppression of E2F transcription factor,  
 CC hypophosphorylation of cellular pbb, or in vitro anti-proliferative  
 CC effects. The peptide is useful in preparing a medicament for treating a  
 CC proliferative disorder, e.g., cancer. The present sequence represents a  
 CC p21-derived peptide of the invention.

XX Sequence 5 AA;

Query Match 95.2%; Score 20; DB 9; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 2e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
 DB 1 RRLN 4  
 ||||

RESULT 16  
 ADZ71747  
 ID ADZ71747 standard; peptide; 5 AA.  
 AC ADZ71747;  
 DT 14-JUL-2005 (first entry)  
 XX  
 XX p21-derived peptide #332.  
 DE  
 DE CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;  
 KW neoplasm; cytostatic; pharmaceutical; drug screening.  
 XX  
 XX Synthetic.  
 OS  
 XX WO2005040802-A2.  
 PN  
 XX 06-MAY-2005.  
 PD  
 XX 20-OCT-2004; 2004WO-GB004431.  
 PF  
 XX 20-OCT-2003; 2003GB-00024466.  
 XX  
 XX (CYCL-) CYCLACEL LTD.  
 XX  
 XX Zheleva DI, Fischer PM, McInnes C, Andrews MJ, Chan WC;  
 PI Atkinson GE;  
 DR WPI; 2005-355897/36.  
 XX  
 XX New peptide inhibitors of cyclin dependent kinases derived from the C-  
 FT terminal region of p21, useful in preparing a medicament for treating a  
 FT proliferative disorder such as cancer.  
 PT  
 PT Claim 21; Page 102; 112pp; English.  
 PS  
 XX The invention relates to a peptide or its variant comprising formula: A-  
 CC (B) m -C-(D) n -E, where m or n are each independently 0 or 1; A is a  
 CC natural or unnatural amino acid residue having a side chain comprising at  
 CC least one H-bond acceptor moiety and at least one H-bond donor moiety;  
 CC each of B or D is independently an amino acid residue selected from  
 CC arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,  
 CC isoleucine or alanine; C is a natural or unnatural amino acid residue  
 CC having a branched or unbranched C1-C6 alkylene side chain optionally  
 CC containing a H-bond donor or a H-bond acceptor moiety; and E is a natural

CC or unnatural amino acid residue having an aryl or heteroaryl side chain.  
 CC Also described are: a pharmaceutical composition comprising the peptide  
 CC admixed with a diluent, an excipient or a carrier; an assay for  
 CC identifying candidate substances capable of binding to a cyclin  
 CC associated with a G1 control CDK enzyme and/or inhibiting the enzyme; an  
 CC assay for identifying compounds that interact a cyclin or a cyclin when  
 CC complexed with the physiologically relevant CDK; and a method of using a  
 CC cyclin in a drug screening assay. The assay for identifying candidate  
 CC substances capable of binding to a cyclin associated with a G1 control  
 CC CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact  
 CC a peptide as defined above, the cyclin, the CDK and the candidate  
 CC substance, under conditions where, in the absence of the cyclin/CDK  
 CC interaction, the peptidomimetic would bind to the cyclin; and monitoring  
 CC any change in the expected binding of the peptide and the cyclin. The  
 CC assay for identifying compounds that interact a cyclin or a cyclin when  
 CC complexed with the physiologically relevant CDK comprises: incubating a  
 CC candidate compound and the peptide and a cyclin or cyclin/CDK complex;  
 CC and detecting binding of either the candidate compound or the peptide  
 CC with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay  
 CC comprises use of a three-dimensional model of a cyclin and a candidate  
 CC compound. At least one of the assay components is bound to a solid phase.  
 CC The peptidomimetic is labeled such as to emit a signal when bound to the  
 CC cyclin. The cyclin is labeled such as to emit a signal when bound to the  
 CC peptide. One of the assay components is labeled with a fluorescence  
 CC emitter and the signal is detected using fluorescence polarization  
 CC techniques. Using a cyclin in a drug screening assay comprises: selecting  
 CC a candidate compound by performing rational drug design with a three-  
 CC dimensional model of the cyclin, where the selecting is performed in  
 CC conjunction with computer modeling; contacting the candidate compound for  
 CC with the cyclin; and detecting the binding of the candidate compound for  
 CC the cyclin groove. A potential drug is selected on the basis of its  
 CC having a greater affinity for the cyclin groove than that of the peptide.  
 CC The method of detection comprises monitoring G0 and/or G1/S cell cycle,  
 CC cell cycle-related apoptosis, suppression of E2F transcription factor,  
 CC hypophosphorylation of cellular pbb, or in vitro anti-proliferative  
 CC effects. The peptide is useful in preparing a medicament for treating a  
 CC proliferative disorder, e.g., cancer. The present sequence represents a  
 CC p21-derived peptide of the invention.

XX Sequence 5 AA;

Query Match 95.2%; Score 20; DB 9; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 2e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
 DB 1 RRLN 4  
 ||||

RESULT 17  
 ADZ71748  
 ID ADZ71748 standard; peptide; 5 AA.  
 XX  
 XX ADZ71748;  
 AC  
 XX 14-JUL-2005 (first entry)  
 DT  
 XX p21-derived peptide #333.  
 DE  
 DE CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;  
 KW neoplasm; cytostatic; pharmaceutical; drug screening.  
 XX  
 XX Synthetic.  
 OS  
 XX WO2005040802-A2.  
 PN  
 XX 06-MAY-2005.  
 PD  
 XX 20-OCT-2004; 2004WO-GB004431.  
 PF  
 XX 20-OCT-2003; 2003GB-00024466.  
 XX  
 XX





RESULT 20  
AAM46173  
ID AAM46173 standard; peptide; 7 AA.  
XX  
AC AAM46173;



KW Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;  
 KW immunogenically cross-reactive; cancer; immunogenic cancer cell;  
 KW cytostatic; vaccine; tumour-specific immunogenic response inducer;  
 KW astrocytoma; fibrosarcoma; myxosarcoma; liposarcoma; oligodendroglioma;  
 KW ependymoma; medulloblastoma; primitive neural ectodermal tumour.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PN CA2290722-A1.  
 XX  
 PD 08-JUN-2001.  
 XX  
 XX 08-DEC-1999; 99CA-02290722.  
 XX  
 PF 08-DEC-1999; 99CA-02290722.  
 XX  
 PR 08-DEC-1999; 99CA-02290722.  
 XX  
 XX (NOVO-) NOVOPHARM BIOTECH INC.  
 XX  
 PI Kaplan HA, Maiti PK, Fast DG, Herman W, Dan MD, Lewis KE;  
 PI Entwistle JM, Macdonald GC;  
 XX  
 XX WPI; 2001-425937/46.  
 XX  
 DR Composition useful for treating and diagnosing cancer, comprises stress  
 PT protein-peptide complexes associated with tumor, and isolated antigen-  
 PT binding fragments of an antibody that binds specifically to the complex.  
 PT  
 XX Example 4; Page 104; 154pp; English.  
 XX  
 CC The present invention describes a composition (I) comprising stress  
 CC protein-peptide complexes (SPPC) associated with tumours that is  
 CC specifically immunogenically cross-reactive with cell surface-associated  
 CC SPPCs specific to target cancer (TC). Also described is an isolated  
 CC antigen-binding fragment of an antibody that binds specifically to SPPCs  
 CC or a population of different SPPCs consisting of immunogenic cancer cell  
 CC surface-associated SPPC of TC. (I) has cytostatic activity and can be  
 CC used in vaccine production and as a tumour-specific immunogenic response  
 CC inducer. (I) is useful for treating 71 types of cancers or tumours in a  
 CC subject, such as astrocytoma, fibrosarcoma, myxosarcoma, liposarcoma,  
 CC oligodendroglioma, ependymoma, medulloblastoma, and primitive neural  
 CC ectodermal tumour (PNET). (I) is useful as cancer immunogen including  
 CC vaccines. (I) is useful for diagnostic and palliative use, for detecting  
 CC or imaging cancer cells, and to monitor the course of amelioration of  
 CC malignancy in an individual. AAM43707 to AAM47109 represent peptides  
 CC which are used in the exemplification of the present invention  
 XX  
 SQ Sequence 7 AA;  
 Query Match 95.2%; Score 20; DB 4; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 2e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RRLN 4  
 Db ||||  
 2 RRLN 5  
 RESULT 22  
 AAM44575  
 ID AAM44575 standard; peptide; 7 AA.  
 XX  
 XX AC AAM44575;  
 AC  
 XX 25-OCT-2001 (first entry)  
 DT  
 XX H11 binding site consensus conforming peptide (CCP) #846.  
 DE  
 DE Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;  
 KW immunogenically cross-reactive; cancer; immunogenic cancer cell;  
 KW cytostatic; vaccine; tumour-specific immunogenic response inducer;  
 KW astrocytoma; fibrosarcoma; myxosarcoma; liposarcoma; oligodendroglioma;  
 KW ependymoma; medulloblastoma; primitive neural ectodermal tumour.

XX Homo sapiens.  
 OS Synthetic.  
 XX  
 PN CA2290722-A1.  
 XX  
 PD 08-JUN-2001.  
 XX  
 XX 08-DEC-1999; 99CA-02290722.  
 XX  
 PF 08-DEC-1999; 99CA-02290722.  
 XX  
 PR 08-DEC-1999; 99CA-02290722.  
 XX  
 XX (NOVO-) NOVOPHARM BIOTECH INC.  
 XX  
 PI Kaplan HA, Maiti PK, Fast DG, Herman W, Dan MD, Lewis KE;  
 PI Entwistle JM, Macdonald GC;  
 XX  
 XX WPI; 2001-425937/46.  
 XX  
 DR Composition useful for treating and diagnosing cancer, comprises stress  
 PT protein-peptide complexes associated with tumor, and isolated antigen-  
 PT binding fragments of an antibody that binds specifically to the complex.  
 PT  
 XX Example 4; Page 104; 154pp; English.  
 XX  
 CC The present invention describes a composition (I) comprising stress  
 CC protein-peptide complexes (SPPC) associated with tumours that is  
 CC specifically immunogenically cross-reactive with cell surface-associated  
 CC SPPCs specific to target cancer (TC). Also described is an isolated  
 CC antigen-binding fragment of an antibody that binds specifically to SPPCs  
 CC or a population of different SPPCs consisting of immunogenic cancer cell  
 CC surface-associated SPPC of TC. (I) has cytostatic activity and can be  
 CC used in vaccine production and as a tumour-specific immunogenic response  
 CC inducer. (I) is useful for treating 71 types of cancers or tumours in a  
 CC subject, such as astrocytoma, fibrosarcoma, myxosarcoma, liposarcoma,  
 CC oligodendroglioma, ependymoma, medulloblastoma, and primitive neural  
 CC ectodermal tumour (PNET). (I) is useful as cancer immunogen including  
 CC vaccines. (I) is useful for diagnostic and palliative use, for detecting  
 CC or imaging cancer cells, and to monitor the course of amelioration of  
 CC malignancy in an individual. AAM43707 to AAM47109 represent peptides  
 CC which are used in the exemplification of the present invention  
 XX  
 SQ Sequence 7 AA;  
 Query Match 95.2%; Score 20; DB 4; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 2e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RRLN 4  
 Db ||||  
 2 RRLN 5  
 RESULT 23  
 AAM43976  
 ID AAM43976 standard; peptide; 7 AA.  
 XX  
 XX AC AAM43976;  
 AC  
 XX 25-OCT-2001 (first entry)  
 DT  
 XX H11 binding site consensus conforming peptide (CCP) #247.  
 DE  
 DE Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;  
 KW immunogenically cross-reactive; cancer; immunogenic cancer cell;  
 KW cytostatic; vaccine; tumour-specific immunogenic response inducer;  
 KW astrocytoma; fibrosarcoma; myxosarcoma; liposarcoma; oligodendroglioma;  
 KW ependymoma; medulloblastoma; primitive neural ectodermal tumour.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PN CA2290722-A1.

```
XX PD 08-JUN-2001.
XX PF 08-DEC-1999; 99CA-02290722.
XX FF 08-DEC-1999; 99CA-02290722.
XX PR 08-DEC-1999; 99CA-02290722.
XX PA (NOVO-) NOVOPHARM BIOTECH INC.
XX PI Kaplan HA, Maiti PK, Fast DG, Herman W, Dan MD, Lewis KE;
XX PI Entwistle JM, Macdonald GC;
XX DR WPI; 2001-425937/46.
XX PS Composition useful for treating and diagnosing cancer, comprises stress
XX PT protein-peptide complexes associated with tumor, and isolated antigen-
XX PT binding fragments of an antibody that binds specifically to the complex.
XX PS Example 4; Page 101; 154pp; English.
XX CC The present invention describes a composition (I) comprising stress
XX CC protein-peptide complexes (SPPC) associated with tumours that is
XX CC specifically immunogenically cross-reactive with cell surface-associated
XX CC SPPCs specific to target cancer (TC). Also described is an isolated
XX CC antigen-binding fragment of an antibody that binds specifically to SPPCs
XX CC or a population of different SPPCs consisting of immunogenic cancer cell
XX CC surface-associated SPPC of TC. (I) has cytostatic activity and can be
XX CC used in vaccine production and as a tumour-specific immunogenic response
XX CC inducer. (I) is useful for treating 71 types of cancers or tumours in a
XX CC subject, such as astrocytoma, fibrosarcoma, myxosarcoma, liposarcoma,
XX CC oligodendroglioma, ependymoma, medulloblastoma, and primitive neural
XX CC ectodermal tumour (PNET). (I) is useful as cancer immunogen including
XX CC vaccines. (I) is useful for diagnostic and palliative use, for detecting
XX CC or imaging cancer cells, and to monitor the course of amelioration of
XX CC malignancy in an individual. AAM43707 to AAM47109 represent peptides
XX CC which are used in the exemplification of the present invention
XX SQ Sequence 7 AA;
XX Query Match 95.2%; Score 20; DB 4; Length 7;
XX Best Local Similarity 100.0%; Pred. No. 2e+06;
XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 1 RRLN 4
XX Db |||||
XX 2 RRLN 5
XX RESULT 24
XX ID AAM44570 standard; peptide; 7 AA.
XX AC AAM44570;
XX DT 25-OCT-2001 (first entry)
XX DE H11 binding site consensus conforming peptide (CCP) #841.
XX KW Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;
XX KW immunogenically cross-reactive; cancer; immunogenic cancer cell;
XX KW cytostatic; vaccine; tumour-specific immunogenic response inducer;
XX KW astrocytoma; fibrosarcoma; myxosarcoma; liposarcoma; oligodendroglioma;
XX KW ependymoma; medulloblastoma; primitive neural ectodermal tumour.
XX OS Homo sapiens.
XX OS Synthetic.
XX PN CA2290722-A1.
XX PD 08-JUN-2001.
XX PF 08-DEC-1999; 99CA-02290722.
XX PI (NOVO-) NOVOPHARM BIOTECH INC.
XX PI Kaplan HA, Maiti PK, Fast DG, Herman W, Dan MD, Lewis KE;
```

```
PR 08-DEC-1999; 99CA-02290722.
XX (NOVO-) NOVOPHARM BIOTECH INC.
XX PI Kaplan HA, Maiti PK, Fast DG, Herman W, Dan MD, Lewis KE;
XX PI Entwistle JM, Macdonald GC;
XX DR WPI; 2001-425937/46.
XX PS Composition useful for treating and diagnosing cancer, comprises stress
XX PT protein-peptide complexes associated with tumor, and isolated antigen-
XX PT binding fragments of an antibody that binds specifically to the complex.
XX PS Example 4; Page 104; 154pp; English.
XX CC The present invention describes a composition (I) comprising stress
XX CC protein-peptide complexes (SPPC) associated with tumours that is
XX CC specifically immunogenically cross-reactive with cell surface-associated
XX CC SPPCs specific to target cancer (TC). Also described is an isolated
XX CC antigen-binding fragment of an antibody that binds specifically to SPPCs
XX CC or a population of different SPPCs consisting of immunogenic cancer cell
XX CC surface-associated SPPC of TC. (I) has cytostatic activity and can be
XX CC used in vaccine production and as a tumour-specific immunogenic response
XX CC inducer. (I) is useful for treating 71 types of cancers or tumours in a
XX CC subject, such as astrocytoma, fibrosarcoma, myxosarcoma, liposarcoma,
XX CC oligodendroglioma, ependymoma, medulloblastoma, and primitive neural
XX CC ectodermal tumour (PNET). (I) is useful as cancer immunogen including
XX CC vaccines. (I) is useful for diagnostic and palliative use, for detecting
XX CC or imaging cancer cells, and to monitor the course of amelioration of
XX CC malignancy in an individual. AAM43707 to AAM47109 represent peptides
XX CC which are used in the exemplification of the present invention
XX SQ Sequence 7 AA;
XX Query Match 95.2%; Score 20; DB 4; Length 7;
XX Best Local Similarity 100.0%; Pred. No. 2e+06;
XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 1 RRLN 4
XX Db |||||
XX 2 RRLN 5
XX RESULT 25
XX ID AAM46480 standard; peptide; 7 AA.
XX AC AAM46480;
XX DT 25-OCT-2001 (first entry)
XX DE H11 binding site consensus conforming peptide (CCP) #2751.
XX KW Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;
XX KW immunogenically cross-reactive; cancer; immunogenic cancer cell;
XX KW cytostatic; vaccine; tumour-specific immunogenic response inducer;
XX KW astrocytoma; fibrosarcoma; myxosarcoma; liposarcoma; oligodendroglioma;
XX KW ependymoma; medulloblastoma; primitive neural ectodermal tumour.
XX OS Homo sapiens.
XX OS Synthetic.
XX PN CA2290722-A1.
XX PD 08-JUN-2001.
XX PF 08-DEC-1999; 99CA-02290722.
XX PR 08-DEC-1999; 99CA-02290722.
XX PA (NOVO-) NOVOPHARM BIOTECH INC.
XX PI Kaplan HA, Maiti PK, Fast DG, Herman W, Dan MD, Lewis KE;
```

PI Entwistle JM, Macdonald GC;  
XX WPI; 2001-425937/46.  
XX  
XX Composition useful for treating and diagnosing cancer, comprises stress  
PT protein-peptide complexes associated with tumor, and isolated antigen-  
PT binding fragments of an antibody that binds specifically to the complex.  
XX  
XX Example 4; Page 110; 154pp; English.  
XX  
XX The present invention describes a composition (I) comprising stress  
CC protein-peptide complexes (SPPC) associated with tumors that is  
CC specifically immunogenically cross-reactive with cell surface-associated  
CC SPPCs specific to target cancer (TC). Also described is an isolated  
CC antigen-binding fragment of an antibody that binds specifically to SPPCs  
CC or a population of different SPPCs consisting of immunogenic cancer cell  
CC surface-associated SPCC of TC. (I) has cytostatic activity and can be  
CC used in vaccine production and as a tumour-specific immunogenic response  
CC inducer. (I) is useful for treating 71 types of cancers or tumours in a  
CC subject, such as astrocytoma, fibrosarcoma, myxosarcoma, liposarcoma,  
CC oligodendroglioma, ependymoma, medulloblastoma, and primitive neural  
CC ectodermal tumour (PNET). (I) is useful as cancer immunogen including  
CC vaccines. (I) is useful for diagnostic and palliative use, for detecting  
CC or imaging cancer cells, and to monitor the course of amelioration of  
CC malignancy in an individual. AAM43707 to AAM47109 represent peptides  
CC which are used in the exemplification of the present invention  
XX  
SQ Sequence 7 AA;  
  
Query Match 95.2%; Score 20; DB 4; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2e+06; 0; Indels 0; Gaps 0;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 RRLN 4  
Db 2 RRLN 5  
  
RESULT 26  
AAM44560  
ID AAM44560 standard; peptide; 7 AA.  
XX  
XX AAM44560;  
XX  
XX 25-OCT-2001 (first entry)  
XX  
XX H11 binding site consensus conforming peptide (CCP) #831.  
XX  
XX Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;  
KW immunogenically cross-reactive; cancer; immunogenic cancer cell;  
KW cytostatic; vaccine; tumour-specific immunogenic response inducer;  
KW astrocytoma; fibrosarcoma; myxosarcoma; liposarcoma; oligodendroglioma;  
KW ependymoma; medulloblastoma; primitive neural ectodermal tumour.  
XX  
XX Homo sapiens.  
OS Synthetic.  
XX  
XX CA2290722-A1.  
XX  
XX 08-JUN-2001.  
XX  
XX 08-DEC-1999; 99CA-02290722.  
XX  
XX 08-DEC-1999; 99CA-02290722.  
XX  
XX (NOVO-) NOVOPHARM BIOTECH INC.  
XX  
XX Kaplan HA, Maiti PK, Fast DG, Herman W, Dan MD, Lewis KE;  
PI Entwistle JM, Macdonald GC;  
XX WPI; 2001-425937/46.  
XX  
XX Composition useful for treating and diagnosing cancer, comprises stress

PT protein-peptide complexes associated with tumor, and isolated antigen-  
PT binding fragments of an antibody that binds specifically to the complex.  
XX  
XX Example 4; Page 104; 154pp; English.  
XX  
XX The present invention describes a composition (I) comprising stress  
CC protein-peptide complexes (SPPC) associated with tumors that is  
CC specifically immunogenically cross-reactive with cell surface-associated  
CC SPPCs specific to target cancer (TC). Also described is an isolated  
CC antigen-binding fragment of an antibody that binds specifically to SPPCs  
CC or a population of different SPPCs consisting of immunogenic cancer cell  
CC surface-associated SPCC of TC. (I) has cytostatic activity and can be  
CC used in vaccine production and as a tumour-specific immunogenic response  
CC inducer. (I) is useful for treating 71 types of cancers or tumours in a  
CC subject, such as astrocytoma, fibrosarcoma, myxosarcoma, liposarcoma,  
CC oligodendroglioma, ependymoma, medulloblastoma, and primitive neural  
CC ectodermal tumour (PNET). (I) is useful as cancer immunogen including  
CC vaccines. (I) is useful for diagnostic and palliative use, for detecting  
CC or imaging cancer cells, and to monitor the course of amelioration of  
CC malignancy in an individual. AAM43707 to AAM47109 represent peptides  
CC which are used in the exemplification of the present invention  
XX  
SQ Sequence 7 AA;  
  
Query Match 95.2%; Score 20; DB 4; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2e+06; 0; Indels 0; Gaps 0;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 RRLN 4  
Db 2 RRLN 5  
  
RESULT 27  
AAM45974  
ID AAM45974 standard; peptide; 7 AA.  
XX  
XX AAM45974;  
XX  
XX 25-OCT-2001 (first entry)  
XX  
XX H11 binding site consensus conforming peptide (CCP) #2245.  
XX  
XX Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;  
KW immunogenically cross-reactive; cancer; immunogenic cancer cell;  
KW cytostatic; vaccine; tumour-specific immunogenic response inducer;  
KW astrocytoma; fibrosarcoma; myxosarcoma; liposarcoma; oligodendroglioma;  
KW ependymoma; medulloblastoma; primitive neural ectodermal tumour.  
XX  
XX Homo sapiens.  
OS Synthetic.  
XX  
XX CA2290722-A1.  
XX  
XX 08-JUN-2001.  
XX  
XX 08-DEC-1999; 99CA-02290722.  
XX  
XX 08-DEC-1999; 99CA-02290722.  
XX  
XX (NOVO-) NOVOPHARM BIOTECH INC.  
XX  
XX Kaplan HA, Maiti PK, Fast DG, Herman W, Dan MD, Lewis KE;  
PI Entwistle JM, Macdonald GC;  
XX WPI; 2001-425937/46.  
XX  
XX Composition useful for treating and diagnosing cancer, comprises stress  
PT protein-peptide complexes associated with tumor, and isolated antigen-  
PT binding fragments of an antibody that binds specifically to the complex.  
XX  
XX Example 4; Page 109; 154pp; English.

CC The present invention describes a composition (I) comprising stress  
 CC protein-peptide complexes (SPPC) associated with tumours that is  
 CC specifically immunogenically cross-reactive with cell surface-associated  
 CC SPPCs specific to target cancer (TC). Also described is an isolated  
 CC antigen-binding fragment of an antibody that binds specifically to SPPCs  
 CC or a population of different SPPCs consisting of immunogenic cancer cell  
 CC surface-associated SPPC of TC. (I) has cytostatic activity and can be  
 CC used in vaccine production and as a tumour-specific immunogenic response  
 CC inducer. (I) is useful for treating 71 types of cancers or tumours in a  
 CC subject, such as astrocytoma, fibrosarcoma, myxosarcoma, liposarcoma,  
 CC oligodendroglioma, ependymoma, medulloblastoma, and primitive neural  
 CC ectodermal tumour (PNET). (I) is useful as cancer immunogen including  
 CC vaccines. (I) is useful for diagnostic and palliative use, for detecting  
 CC or imaging cancer cells, and to monitor the course of amelioration of  
 CC malignancy in an individual. AAM43707 to AAM47109 represent peptides  
 CC which are used in the exemplification of the present invention  
 XX  
 SQ Sequence 7 AA;  
 Query Match 95.2%; Score 20; DB 4; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 2e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RRLN 4  
 Db 2 RRLN 5  
 RESULT 28  
 ID AAM46528 standard; peptide; 7 AA.  
 AC AAM46528;  
 XX  
 DT 25-OCT-2001 (first entry)  
 XX  
 DE H11 binding site consensus conforming peptide (CCP) #2799.  
 XX  
 KW Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;  
 KW immunogenically cross-reactive; cancer; immunogenic cancer cell;  
 KW cytostatic; vaccine; tumour-specific immunogenic response inducer;  
 KW astrocytoma; fibrosarcoma; myxosarcoma; liposarcoma; oligodendroglioma;  
 KW ependymoma; medulloblastoma; primitive neural ectodermal tumour.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PN CA2290722-A1.  
 XX  
 PD 08-JUN-2001.  
 XX  
 PF 08-DEC-1999; 99CA-02290722.  
 XX  
 PR 08-DEC-1999; 99CA-02290722.  
 XX  
 PA (NOVO-) NOVOPHARM BIOTECH INC.  
 XX  
 PI Kaplan HA, Maiti PK, Fast DG, Herman W, Dan MD, Lewis KE;  
 PI Entwistle JM, Macdonald GC;  
 XX  
 DR WPI; 2001-425937/46.  
 XX  
 XX Composition useful for treating and diagnosing cancer, comprises stress  
 PT protein-peptide complexes associated with tumor, and isolated antigen-  
 PT binding fragments of an antibody that binds specifically to the complex.  
 XX  
 PS Example 4; Page 111; 154pp; English.  
 XX  
 XX The present invention describes a composition (I) comprising stress  
 CC protein-peptide complexes (SPPC) associated with tumours that is  
 CC specifically immunogenically cross-reactive with cell surface-associated  
 CC SPPCs specific to target cancer (TC). Also described is an isolated  
 CC antigen-binding fragment of an antibody that binds specifically to SPPCs  
 CC or a population of different SPPCs consisting of immunogenic cancer cell  
 CC surface-associated SPPC of TC. (I) has cytostatic activity and can be  
 CC used in vaccine production and as a tumour-specific immunogenic response  
 CC inducer. (I) is useful for treating 71 types of cancers or tumours in a  
 CC subject, such as astrocytoma, fibrosarcoma, myxosarcoma, liposarcoma,  
 CC oligodendroglioma, ependymoma, medulloblastoma, and primitive neural  
 CC ectodermal tumour (PNET). (I) is useful as cancer immunogen including  
 CC vaccines. (I) is useful for diagnostic and palliative use, for detecting  
 CC or imaging cancer cells, and to monitor the course of amelioration of  
 CC malignancy in an individual. AAM43707 to AAM47109 represent peptides  
 CC which are used in the exemplification of the present invention  
 XX

CC or a population of different SPPCs consisting of immunogenic cancer cell  
 CC surface-associated SPPC of TC. (I) has cytostatic activity and can be  
 CC used in vaccine production and as a tumour-specific immunogenic response  
 CC inducer. (I) is useful for treating 71 types of cancers or tumours in a  
 CC subject, such as astrocytoma, fibrosarcoma, myxosarcoma, liposarcoma,  
 CC oligodendroglioma, ependymoma, medulloblastoma, and primitive neural  
 CC ectodermal tumour (PNET). (I) is useful as cancer immunogen including  
 CC vaccines. (I) is useful for diagnostic and palliative use, for detecting  
 CC or imaging cancer cells, and to monitor the course of amelioration of  
 CC malignancy in an individual. AAM43707 to AAM47109 represent peptides  
 CC which are used in the exemplification of the present invention  
 XX

SQ Sequence 7 AA;

Query Match 95.2%; Score 20; DB 4; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 2e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4

Db 2 RRLN 5

RESULT 29

AAM46485

ID AAM46485 standard; peptide; 7 AA.

AC AAM46485;

DT 25-OCT-2001 (first entry)

DE H11 binding site consensus conforming peptide (CCP) #2756.

KW Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;  
 KW immunogenically cross-reactive; cancer; immunogenic cancer cell;  
 KW cytostatic; vaccine; tumour-specific immunogenic response inducer;  
 KW astrocytoma; fibrosarcoma; myxosarcoma; liposarcoma; oligodendroglioma;  
 KW ependymoma; medulloblastoma; primitive neural ectodermal tumour.

OS Homo sapiens.

OS Synthetic.

PN CA2290722-A1.

XX 08-JUN-2001.

XX 08-DEC-1999; 99CA-02290722.

XX 08-DEC-1999; 99CA-02290722.

XX (NOVO-) NOVOPHARM BIOTECH INC.

XX Kaplan HA, Maiti PK, Fast DG, Herman W, Dan MD, Lewis KE;

XX Entwistle JM, Macdonald GC;

XX WPI; 2001-425937/46.

XX Composition useful for treating and diagnosing cancer, comprises stress  
 PT protein-peptide complexes associated with tumor, and isolated antigen-  
 PT binding fragments of an antibody that binds specifically to the complex.

PS Example 4; Page 110; 154pp; English.

XX The present invention describes a composition (I) comprising stress

CC protein-peptide complexes (SPPC) associated with tumours that is  
 CC specifically immunogenically cross-reactive with cell surface-associated  
 CC SPPCs specific to target cancer (TC). Also described is an isolated  
 CC antigen-binding fragment of an antibody that binds specifically to SPPCs  
 CC or a population of different SPPCs consisting of immunogenic cancer cell  
 CC surface-associated SPPC of TC. (I) has cytostatic activity and can be  
 CC used in vaccine production and as a tumour-specific immunogenic response  
 CC inducer. (I) is useful for treating 71 types of cancers or tumours in a  
 CC subject, such as astrocytoma, fibrosarcoma, myxosarcoma, liposarcoma,

CC oligodendroglioma, ependymoma, medulloblastoma, and primitive neural  
CC ectodermal tumour (PNET). (I) is useful as cancer immunogen including  
CC vaccines. (I) is useful for diagnostic and palliative use, for detecting  
CC or imaging cancer cells, and to monitor the course of amelioration of  
CC malignancy in an individual. AAM43707 to AAM47109 represent peptides  
CC which are used in the exemplification of the present invention  
XX  
SQ Sequence 7 AA;

Query Match 95.2%; Score 20; DB 4; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
Db 2 RRLN 5  
|||||

RESULT 30  
AAM45665  
ID AAM45665 standard; peptide; 7 AA.  
XX  
AC AAM45665;  
XX  
XX  
DT 25-OCT-2001 (first entry)  
XX  
DE H11 binding site consensus conforming peptide (CCP) #1936.  
XX  
KW Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;  
KW immunogenically cross-reactive; cancer; immunogenic cancer cell;  
KW cytostatic; vaccine; tumour-specific immunogenic response inducer;  
KW astrocytoma; fibrosarcoma; myxosarcoma; liposarcoma; oligodendroglioma;  
KW ependymoma; medulloblastoma; primitive neural ectodermal tumour.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
XX  
PN CA2290722-A1.  
XX  
PD 08-JUN-2001.  
XX  
PF 08-DEC-1999; 99CA-02290722.  
XX  
PR 08-DEC-1999; 99CA-02290722.  
XX  
XX (NOVO-) NOVOPHARM BIOTECH INC.  
XX  
PI Kaplan HA, Maiti PK, Fast DG, Herman W, Dan MD, Lewis KE;  
PI Entwistle JM, Macdonald GC;  
XX  
DR WPI; 2001-425937/46.  
XX  
XX  
PT Composition useful for treating and diagnosing cancer, comprises stress  
PT protein-peptide complexes associated with tumor, and isolated antigen-  
PT binding fragments of an antibody that binds specifically to the complex.  
XX  
PS Example 4; Page 107; 154pp; English.  
XX  
XX The present invention describes a composition (I) comprising stress  
CC protein-peptide complexes (SPPC) associated with tumors that is  
CC specifically immunogenically cross-reactive with cell surface-associated  
CC SPPCs specific to target cancer (TC). Also described is an isolated  
CC antigen-binding fragment of an antibody that binds specifically to SPPCs  
CC or a population of different SPPCs consisting of immunogenic cancer cell  
CC surface-associated SPPC of TC. (I) has cytostatic activity and can be  
CC used in vaccine production and as a tumour-specific immunogenic response  
CC inducer. (I) is useful for treating 71 types of cancers or tumours in a  
CC subject, such as astrocytoma, fibrosarcoma, myxosarcoma, liposarcoma,  
CC oligodendroglioma, ependymoma, medulloblastoma, and primitive neural  
CC ectodermal tumour (PNET). (I) is useful as cancer immunogen including  
CC vaccines. (I) is useful for diagnostic and palliative use, for detecting  
CC or imaging cancer cells, and to monitor the course of amelioration of  
CC malignancy in an individual. AAM43707 to AAM47109 represent peptides

CC which are used in the exemplification of the present invention  
XX  
SQ Sequence 7 AA;

Query Match 95.2%; Score 20; DB 4; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
Db 2 RRLN 5  
|||||

RESULT 31  
AAM46523  
ID AAM46523 standard; peptide; 7 AA.  
XX  
AC AAM46523;  
XX  
XX  
DT 25-OCT-2001 (first entry)  
XX  
DE H11 binding site consensus conforming peptide (CCP) #2794.  
XX  
KW Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;  
KW immunogenically cross-reactive; cancer; immunogenic cancer cell;  
KW cytostatic; vaccine; tumour-specific immunogenic response inducer;  
KW astrocytoma; fibrosarcoma; myxosarcoma; liposarcoma; oligodendroglioma;  
KW ependymoma; medulloblastoma; primitive neural ectodermal tumour.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
XX  
PN CA2290722-A1.  
XX  
PD 08-JUN-2001.  
XX  
PF 08-DEC-1999; 99CA-02290722.  
XX  
PR 08-DEC-1999; 99CA-02290722.  
XX  
XX (NOVO-) NOVOPHARM BIOTECH INC.  
XX  
PI Kaplan HA, Maiti PK, Fast DG, Herman W, Dan MD, Lewis KE;  
PI Entwistle JM, Macdonald GC;  
XX  
DR WPI; 2001-425937/46.  
XX  
XX  
PT Composition useful for treating and diagnosing cancer, comprises stress  
PT protein-peptide complexes associated with tumor, and isolated antigen-  
PT binding fragments of an antibody that binds specifically to the complex.  
XX  
PS Example 4; Page 111; 154pp; English.  
XX  
XX The present invention describes a composition (I) comprising stress  
CC protein-peptide complexes (SPPC) associated with tumors that is  
CC specifically immunogenically cross-reactive with cell surface-associated  
CC SPPCs specific to target cancer (TC). Also described is an isolated  
CC antigen-binding fragment of an antibody that binds specifically to SPPCs  
CC or a population of different SPPCs consisting of immunogenic cancer cell  
CC surface-associated SPPC of TC. (I) has cytostatic activity and can be  
CC used in vaccine production and as a tumour-specific immunogenic response  
CC inducer. (I) is useful for treating 71 types of cancers or tumours in a  
CC subject, such as astrocytoma, fibrosarcoma, myxosarcoma, liposarcoma,  
CC oligodendroglioma, ependymoma, medulloblastoma, and primitive neural  
CC ectodermal tumour (PNET). (I) is useful as cancer immunogen including  
CC vaccines. (I) is useful for diagnostic and palliative use, for detecting  
CC or imaging cancer cells, and to monitor the course of amelioration of  
CC malignancy in an individual. AAM43707 to AAM47109 represent peptides  
XX  
SQ Sequence 7 AA;

Query Match 95.2%; Score 20; DB 4; Length 7;

Fri Jan 27 10:08:32 2006

us-10-771-242-295.rag

Best Local Similarity 100.0%; Pred. No. 2e+06; Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4  
|  
|  
|  
|  
Db 2 RRLN 5

RESULT 32  
AAM46526  
ID AAM46526 standard; peptide; 7 AA.

AC AAM46526;

XX 25-OCT-2001 (first entry)

XX H11 binding site consensus conforming peptide (CCP) #2797.

XX Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;  
KW immunogenically cross-reactive; cancer; immunogenic cancer cell;  
KW cytostatic; vaccine; tumour-specific immunogenic response inducer;  
KW astrocytoma; fibrosarcoma; myxosarcoma; liposarcoma; oligodendroglioma;  
KW ependymoma; medulloblastoma; primitive neural ectodermal tumour.

OS Homo sapiens.  
OS Synthetic.

PN CA2290722-A1.

PD 08-JUN-2001.

PF 08-DEC-1999; 99CA-02290722.

PR 08-DEC-1999; 99CA-02290722.

XX (NOVO-) NOVOPHARM BIOFTECH INC.

XX Kaplan HA, Maiti PK, Fast DG, Herman W, Dan MD, Lewis KE;  
PI Entwistle JM, Macdonald GC;

XX WPI; 2001-425937/46.

XX Composition useful for treating and diagnosing cancer, comprises stress  
PT protein-peptide complexes associated with tumor, and isolated antigen-  
PT binding fragments of an antibody that binds specifically to the complex.  
XX Example 4; Page 111; 154pp; English.

XX The present invention describes a composition (I) comprising stress  
CC protein-peptide complexes (SPPC) associated with tumours that is  
CC specifically immunogenically cross-reactive with cell surface-associated  
CC SPPCs specific to target cancer (TC). Also described is an isolated  
CC antigen-binding fragment of an antibody that binds specifically to SPPCs  
CC or a population of different SPPCs consisting of immunogenic cancer cell  
CC surface-associated SPPC of TC. (I) has cytostatic activity and can be  
CC used in vaccine production and as a tumour-specific immunogenic response  
CC inducer. (I) is useful for treating 71 types of cancers or tumours in a  
CC subject, such as astrocytoma, fibrosarcoma, myxosarcoma, liposarcoma,  
CC oligodendroglioma, ependymoma, medulloblastoma, and primitive neural  
CC ectodermal tumour (PNET). (I) is useful as cancer immunogen including  
CC vaccines. (I) is useful for diagnostic and palliative use, for detecting  
CC or imaging cancer cells, and to monitor the course of amelioration of  
CC malignancy in an individual. AAM43707 to AAM47109 represent peptides  
CC which are used in the exemplification of the present invention

XX Sequence 7 AA;

Query Match 95.2%; Score 20; DB 4; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4  
|  
|  
|  
|

Db 2 RRLN 5

RESULT 33

AAM57007

ID AAM57007 standard; peptide; 8 AA.

XX AAM57007;

XX 28-JUL-1998 (first entry)

XX Enzyme inhibitor peptide SEQ ID NO:208.

XX Enzyme inhibitor; t-PA; u-PA; chymotrypsin; serine protease; active;  
KW latent; substrate subtraction phage display peptide library;  
KW identification; kinase; phosphatase; serpin.

XX Homo sapiens.

XX WO9747314-A1.

XX 18-DEC-1997.

XX 10-JUN-1997; 97WO-US0009760.

XX 10-JUN-1996; 96US-0019495P.

XX (SCRI) SCRIPPS RES INST.

XX Madison EL, Ke S;

XX WPI; 1998-062746/06.

XX Substrate subtraction phage display peptide libraries - used to  
PT distinguish between active and latent forms of enzyme, e.g. serine  
PT protease.

XX Claim 25; Page 111; 138pp; English.

XX The present sequence represents an enzyme inhibitor peptide used in the  
CC method of the invention to distinguish between t-PA and u-PA. The present  
CC invention describes a substrate subtraction library for the  
CC identification of peptide substrates selective between a first enzyme  
CC (E1) and a second enzyme (E2), comprising a collection of different  
CC peptides, substantially lacking peptides that are effective substrates  
CC for E1. Also described are: (1) a method (M1) for identifying peptide  
CC substrates selective between a first enzyme (E1) and a second enzyme (E2)  
CC; (2) a compound comprising the amino acid sequence of a peptide  
CC identified by M1; (3) a polypeptide for use as an enzyme inhibitor  
CC comprising one of 237 amino acid sequences (see AAM56801 to AAM56947, and  
CC AAM56949 to AAM57038); (4) a recombinant DNA vector comprising DNA (1)  
CC encoding a protease inhibitor including the sequence identified by the M1  
CC; (5) a prokaryotic or eukaryotic cell containing the vector of (4); (6)  
CC an antibody (Ab) immunoreactive with at least one of the peptides  
CC identified by M1; and (7) a diagnostic assay for distinguishing between  
CC active and latent forms of protease inhibitors, that uses (Ab). The  
CC library and method are used for distinguishing between active and latent  
CC forms of enzyme inhibitors, e.g. proteases, kinases and phosphatases.  
CC (Ab) are used for affinity purification of recombinant peptides and in  
CC the identification of naturally occurring protease inhibitors. Enzyme-  
CC inhibiting peptides identified can be used to treat a serpin deficiency  
CC or a disorder of serine proteases

XX Sequence 8 AA;

Query Match 95.2%; Score 20; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4  
|  
|  
|  
|  
Db 4 RRLN 7

RESULT 34  
AAU05744  
ID AAU05744 standard; protein; 8 AA.  
XX AC AAU05744;  
XX AC  
XX AC  
DT 21-NOV-2001 (first entry)  
XX  
DE p21 C-terminus derived peptide #103.  
XX  
KW Human; p21WAF1; cyclin dependent protein kinase; CDK2; cyclin A;  
KW inhibitor; proliferative disorder; cancer; leukaemia; drug screening.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 1 /note= "The N-terminus is hydrogenated"  
FT Modified-site 8 /label= OTHER  
FT FT /note= "parafluorophenylalanine. C-terminal amide"  
XX  
PN WC200140142-A2.  
XX  
PD 07-JUN-2001.  
XX  
PF 29-NOV-2000; 2000WO-GB004550.  
XX  
PR 30-NOV-1999; 99GB-00028323.  
XX  
PA (CYCL-) CYCLACEL LTD.  
XX  
PI Zheleva DI, Fischer PM, McInnes C, Andrews MJ, Chan WC;  
PI Atkinson GE;  
XX  
DR WPI; 2001-488493/53.  
XX  
XX New p21 derived peptides and their variants, particularly useful as  
PT selective inhibitors of CDK2/cyclin interaction for treating  
PT proliferative disorders e.g. cancers and leukemias, and in assays for  
PT identifying CDK/cyclin inhibitors.  
XX  
XX Claim 25; Page 89; 102pp; English.  
XX  
CC The invention relates to peptide and their variants derived from p21WAF1,  
CC which are inhibitors of CDK2 activity by binding to G1 and S phase  
CC specific cyclins which activate CDK2; selective inhibitors of CDK2/cyclin  
CC complexes, particularly CDK2/cyclin A or E complexes. The variants of the  
CC peptide may have further amino acids at either end or have up to 7 amino  
CC acids deleted, provided the motif XLXF is retained. The peptides are  
CC specific regions of p21WAF1 that bind to G1 and S phase specific cyclins,  
CC preferably cyclins which activate CDK2. One of the peptides corresponds  
CC to p21(149-159). The peptides are used for treating proliferative  
CC disorders, e.g. cancers and leukemias. The peptides are also for  
CC identifying substances which interfere with protein-protein interactions  
CC involving cyclins (i.e. cyclin A, E or D), especially CDK/cyclin  
CC interactions, and which are capable of inhibiting CDK2 and/or CDK4  
CC activity. P21 peptides other than p21(149-159) competitively inhibit the  
CC binding of peptide p21(149-159) to cyclin and may be used to identify  
CC substances that bind to, or inhibit peptide- cyclin interactions.  
CC Substances for screening in the assays include antibody products specific  
CC for p21 or cyclin binding regions, combinatorial libraries and single  
CC compound collections. The present sequence is a peptide derived from the  
CC C-terminus of p21  
XX  
SQ Sequence 8 AA;

Query Match 95.2%; Score 20; DB 4; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
Db ||||  
4 RRLN 7

RESULT 35  
ADZ71972  
ID ADZ71972 standard; peptide; 8 AA.  
XX AC ADZ71972;  
XX  
DT 14-JUL-2005 (first entry)  
XX  
DE p21-derived peptide #557.  
XX  
KW CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;  
KW neoplasm; cytostatic; pharmaceutical; drug screening.  
XX  
OS Synthetic.  
XX  
PN WC2005040802-A2.  
XX  
PD 06-MAY-2005.  
XX  
PF 20-OCT-2004; 2004WO-GB004431.  
XX  
PR 20-OCT-2003; 2003GB-00024466.  
PR 02-FEB-2004; 2004US-00771242.  
XX  
PA (CYCL-) CYCLACEL LTD.  
XX  
PI Zheleva DI, Fischer PM, McInnes C, Andrews MJ, Chan WC;  
PI Atkinson GE;  
XX  
DR WPI; 2005-355897/36.  
XX  
PT New peptide inhibitors of cyclin dependent kinases derived from the C-  
PT terminal region of p21, useful in preparing a medicament for treating a  
PT proliferative disorder such as cancer.  
XX  
XX Example 25; Page 81; 112pp; English.

CC The invention relates to a peptide or its variant comprising formula: A-  
CC (B) m -C-(D) n -E, where m or n are each independently 0 or 1; A is a  
CC natural or unnatural amino acid residue having a side chain comprising at  
CC least one H-bond acceptor moiety and at least one H-bond donor moiety;  
CC each of B or D is independently an amino acid residue selected from  
CC arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,  
CC isoleucine or alanine; C is a natural or unnatural amino acid residue  
CC having a branched or unbranched C 1 -C 6 alkylene side chain optionally  
CC containing a H-bond donor or a H-bond acceptor moiety; and E is a natural  
CC or unnatural amino acid residue having an aryl or heteroaryl side chain.  
CC Also described are: a pharmaceutical composition comprising the peptide  
CC admixed with a diluent, an excipient or a carrier; an assay for  
CC identifying candidate substances capable of binding to a cyclin  
CC associated with a G1 control CDK enzyme and/or inhibiting the enzyme;  
CC assay for identifying compounds that interact a cyclin or a cyclin when  
CC complexed with the physiologically relevant CDK; and a method of using a  
CC cyclin in a drug screening assay. The assay for identifying candidate  
CC substances capable of binding to a cyclin associated with a G1 control  
CC CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact  
CC a peptide as defined above, the cyclin, the CDK and the candidate  
CC substance, under conditions where, in the absence of the candidate  
CC substance being an inhibitor of interaction of the cyclin/CDK  
CC interaction, the peptidomimetic would bind to the cyclin; and monitoring  
CC any change in the expected binding of the peptide and the cyclin. The  
CC assay for identifying compounds that interact a cyclin or a cyclin when  
CC complexed with the physiologically relevant CDK comprises: incubating a  
CC candidate compound and the peptide and a cyclin or cyclin/CDK complex;  
CC and detecting binding of either the candidate compound or the peptide  
CC with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay  
CC comprises use of a three-dimensional model of a cyclin and a candidate  
CC compound. At least one of the assay components is bound to a solid phase.

CC The peptidomimetic is labeled such as to emit a signal when bound to the  
 CC cyclin. The cyclin is labeled such as to emit a signal when bound to the  
 CC peptide. One of the assay components is labeled with a fluorescence  
 CC emitter and the signal is detected using fluorescence polarization  
 CC techniques. Using a cyclin in a drug screening assay comprises: selecting  
 CC a candidate compound by performing rational drug design with a three-  
 CC dimensional model of the cyclin, where the selecting is performed in  
 CC conjunction with computer modeling; contacting the candidate compound  
 CC with the cyclin; and detecting the binding of the candidate compound for  
 CC the cyclin groove. A potential drug is selected on the basis of its  
 CC having a greater affinity for the cyclin groove than that of the peptide.  
 CC The method of detection comprises monitoring G0 and/or G1/S cell cycle,  
 CC cell cycle-related apoptosis, suppression of E2F transcription factor,  
 CC hypophosphorylation of cellular pRb, or in vitro anti-proliferative  
 CC effects. The peptide is useful in preparing a medicament for treating a  
 CC proliferative disorder, e.g., cancer. The present sequence represents a  
 CC p21-derived peptide of the invention.

XX Sequence 8 AA;

Query Match 95.2%; Score 20; DB 9; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 2e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
 ||||  
 Db 4 RRLN 7

RESULT 36

ADZ71569

ID ADZ71569 standard; peptide; 8 AA.

AC ADZ71569;

DT 14-JUL-2005 (first entry)

XX p21-derived peptide #154.

XX CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;

XX neoplasm; cytostatic; pharmaceutical; drug screening.

XX Synthetic.

XX WO2005040802-A2.

XX 06-MAY-2005.

XX 20-OCT-2004; 2004WO-GB004431.

XX 20-OCT-2003; 2003GB-00024466.

XX 02-FEB-2004; 2004US-00771242.

XX (CYCL-) CYCLACEL LTD.

XX Zheleva DI, Fischer PM, McInnes C, Andrews MJ, Chan WC;

XX PI Atkinson GE;

XX WPI; 2005-355897/36.

XX New peptide inhibitors of cyclin dependent kinases derived from the C-

XX terminal region of p21, useful in preparing a medicament for treating a

XX proliferative disorder such as cancer.

XX Disclosure; Page 14; 112pp; English.

XX The invention relates to a peptide or its variant comprising formula: A-  
 CC (B) m -C-(D) n -E, where m or n are each independently 0 or 1; A is a  
 CC natural or unnatural amino acid residue having a side chain comprising at  
 CC least one H-bond acceptor moiety and at least one H-bond donor moiety;  
 CC each of B or D is independently an amino acid residue selected from  
 CC arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,  
 CC isoleucine or alanine; C is a natural or unnatural amino acid residue

CC having a branched or unbranched C 1 -C 6 alkylene side chain optionally  
 CC containing a H-bond donor or a H-bond acceptor moiety; and E is a natural  
 CC or unnatural amino acid residue having an aryl or heteroaryl side chain.  
 CC Also described are: a pharmaceutical composition comprising the peptide  
 CC admixed with a diluent, an excipient or a carrier; an assay for  
 CC identifying candidate substances capable of binding to a cyclin  
 CC associated with a G1 control CDK enzyme and/or inhibiting the enzyme; an  
 CC assay for identifying compounds that interact a cyclin or a cyclin when  
 CC complexed with the physiologically relevant CDK; and a method of using a  
 CC cyclin in a drug screening assay. The assay for identifying candidate  
 CC substances capable of binding to a cyclin associated with a G1 control  
 CC CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact  
 CC a peptide as defined above, the cyclin, the CDK and the candidate  
 CC substance, under conditions where, in the absence of the candidate  
 CC substance being an inhibitor of interaction of the cyclin/CDK  
 CC interaction, the peptidomimetic would bind to the cyclin; and monitoring  
 CC any change in the expected binding of the peptide and the cyclin. The  
 CC assay for identifying compounds that interact a cyclin or a cyclin when  
 CC complexed with the physiologically relevant CDK comprises: incubating a  
 CC candidate compound and the peptide and a cyclin or cyclin/CDK complex;  
 CC and detecting binding of either the candidate compound or the peptide  
 CC with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay  
 CC comprises use of a three-dimensional model of a cyclin and a candidate  
 CC compound. At least one of the assay components is bound to a solid phase.  
 CC The peptidomimetic is labeled such as to emit a signal when bound to the  
 CC cyclin. The cyclin is labeled such as to emit a signal when bound to the  
 CC peptide. One of the assay components is labeled with a fluorescence  
 CC emitter and the signal is detected using fluorescence polarization  
 CC techniques. Using a cyclin in a drug screening assay comprises: selecting  
 CC a candidate compound by performing rational drug design with a three-  
 CC dimensional model of the cyclin, where the selecting is performed in  
 CC conjunction with computer modeling; contacting the candidate compound for  
 CC with the cyclin; and detecting the binding of the candidate compound for  
 CC the cyclin groove. A potential drug is selected on the basis of its  
 CC having a greater affinity for the cyclin groove than that of the peptide.  
 CC The method of detection comprises monitoring G0 and/or G1/S cell cycle,  
 CC cell cycle-related apoptosis, suppression of E2F transcription factor,  
 CC hypophosphorylation of cellular pRb, or in vitro anti-proliferative  
 CC effects. The peptide is useful in preparing a medicament for treating a  
 CC proliferative disorder, e.g., cancer. The present sequence represents a  
 CC p21-derived peptide of the invention.

XX Sequence 8 AA;

Query Match 95.2%; Score 20; DB 9; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 2e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
 ||||  
 Db 4 RRLN 7

RESULT 37

AAE18742

ID AAE18742 standard; peptide; 9 AA.

XX AAE18742;

XX 17-MAY-2002 (first entry)

XX Human leucocyte antigen (HLA) class I epitope #6.

XX Human; lung tumour associated antigen; CASB761; vaccine; lung cancer;  
 CC immunotherapeutic; lung preneoplastic lesion; autoimmune disease;  
 CC gene therapy; cytostatic; immunosuppressive; human leucocyte antigen;  
 CC HLA; epitope.

XX Homo sapiens.

XX WO200206338-A1.

XX 24-JAN-2002.





XX  
PS Claim 9; Page 68; 92pp; English.  
XX  
CC The invention relates to vaccines comprising lung tumour associated  
CC antigen referred as CASB761 and its polynucleotide. CASB761 and its DNA  
CC are useful in the manufacture of a vaccine for immunotherapeutically  
CC treating a patient suffering from or susceptible to lung cancer, lung  
CC preneoplastic lesions or other related conditions. Vaccines of the  
CC invention are useful in medicine, for treating cancer, particularly lung  
CC cancer, autoimmune diseases and other related conditions. CASB761  
CC polynucleotides and their proteins are useful as diagnostic reagents, to  
CC diagnose different forms and states of cancer, in staging cancerous  
CC disorder and grading the nature of the cancerous tissue. An antibody  
CC immunospecific for CASB761 is useful to isolate and to identify clones  
CC expressing CASB761 protein or to purify the polypeptide by affinity  
CC chromatography and to treat or prevent, particularly lung cancer,  
CC autoimmune disease and related conditions. CASB761 DNA is used in gene  
CC therapy. The present sequence is human leucocyte antigen (HLA) class II  
CC epitope. This sequence is used to incorporate an epitope of CASB761  
XX  
XX Sequence 9 AA;  
SQ  
Query Match 95.2%; Score 20; DB 5; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2e+06; Indels 0; Gaps 0;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RRLN 4  
DB 6 RRLN 9  
RESULT 40  
AAB80739  
ID AAB80739 standard; peptide; 11 AA.  
XX  
AC AAB80739;  
XX  
DT 02-MAY-2001 (first entry)  
XX  
XX Semenogelin II hK2 cleavage site #6.  
DE  
XX  
XX Cleavage; kallikrein 2; hK2; prodrug.  
XX  
XX Synthetic.  
OS  
XX WO200109165-A2.  
PN  
XX  
XX 08-FEB-2001.  
PD  
XX  
XX 28-JUL-2000; 2000WO-US040496.  
PP  
XX  
XX 29-JUL-1999; 99US-0146316P.  
PR  
XX  
XX (UYJO ) UNIV JOHNS HOPKINS.  
PA  
XX  
XX Denmeade SR, Isaacs JT, Lilja H, Christensen SB;  
PI  
XX WPI; 2001-191450/19.  
DR  
XX  
XX New peptides containing cleavage sites specifically cleaved by human  
PT kallikrein 2, useful for producing prodrugs which treat hK2-producing  
PT cell proliferative disorders without exhibiting non-specific toxicity.  
PT  
XX  
XX Disclosure; Fig 1; 38pp; English.  
PS  
XX  
XX The present invention relates to a peptide comprising an amino acid  
CC sequence having a cleavage site specific for an enzyme having a  
CC proteolytic activity of human kallikrein 2 (hK2), and which is up to 20  
CC amino acids in length. The invention is useful for producing a prodrug  
CC which involves linking a drug which contains a primary amine to the  
CC peptide, in which the linking of the peptide to the drug inhibits the  
CC therapeutic activity of the drug

XX  
SQ Sequence 11 AA;  
XX  
Query Match 95.2%; Score 20; DB 4; Length 11;  
Best Local Similarity 100.0%; Pred. No. 2.9e+02; Indels 0; Gaps 0;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RRLN 4  
DB 7 RRLN 10  
RESULT 41  
ABW01200  
ID ABW01200 standard; peptide; 11 AA.  
XX  
AC ABW01200;  
XX  
DT 15-JAN-2004 (first entry)  
XX  
XX Saccharomyces cerevisiae consensus peptide #13.  
DE  
XX Modulator of translation termination; MTT1; helicase B; antiviral;  
XX therapy; HCSB; nonsense mutation; yeast.  
KW  
XX Saccharomyces cerevisiae.  
OS  
XX US6630294-B1.  
FN  
XX  
XX 07-OCT-2003.  
PD  
XX  
XX 22-JUL-1999; 99US-00359268.  
PF  
XX  
XX 22-JUL-1998; 98US-0093685P.  
PR  
XX  
XX (UYNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.  
PA  
XX  
XX Peltz S, Czaplinski K, Dinman JD;  
PI  
XX WPI; 2003-810549/76.  
DR  
XX  
XX Identifying an agent that increases nonsense suppression, for antiviral  
PT therapy, by contacting modulator of translation termination (Mtt1) in  
PT Saccharomyces cerevisiae with a test agent, and detecting specific  
PT binding to Mtt1.  
XX  
XX Disclosure; Col 49; Opp; English.  
PS  
XX  
XX The invention relates to a method of identifying an agent that increases  
CC nonsense suppression, by contacting modulator of translation termination  
CC (Mtt1) also referred to as helicase B (HCSB) in Saccharomyces cerevisiae.  
CC The method is useful for identifying compositions or agents which  
CC increase nonsense suppression. The invention may also be used for  
CC antiviral therapy and for suppression of pathological nonsense mutations.  
CC The present sequence is Saccharomyces cerevisiae consensus peptide  
XX  
XX Sequence 11 AA;  
SQ  
Query Match 95.2%; Score 20; DB 7; Length 11;  
Best Local Similarity 100.0%; Pred. No. 2.9e+02; Indels 0; Gaps 0;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RRLN 4  
DB 1 RRLN 4  
RESULT 42  
ADP44126  
ID ADP44126 standard; peptide; 11 AA.  
XX  
XX ADP44126;  
AC  
XX

DT 18-NOV-2004 (first entry)  
 DE Yeast translation termination modulation protein related peptide #15.  
 XX gene therapy; translation termination; RNA helicase; MTT1;  
 KW frameshift frequency; aberrant transcript degradation;  
 KW peptidyl transferase modulation; beta-thalassemia; beta-globin;  
 KW Duchenne/Becker Muscular Dystrophy; Haemophilia A; Haemophilia B;  
 KW Von Willebrand Disease; Osteogenesis Imperfecta; Breast cancer;  
 KW Ovarian Cancer; Wilms Tumour; Hirschsprung disease; Cystic fibrosis;  
 KW Kidney Stone; Familial hypercholesterolaemia; Retinitis Pigmentosa;  
 KW Neurofibromatosis; Retinoblastoma; ATM; Costmann Disease; yeast.  
 XX  
 XX Saccharomyces cerevisiae.  
 OS US2004115787-A1.  
 PN 17-JUN-2004.  
 XX  
 XX 28-AUG-2003; 2003US-00652334.  
 PF 22-JUL-1998; 98US-0093685P.  
 PR 22-JUL-1999; 99US-00359268.  
 XX  
 XX (PELTZ/) PELTZ S.  
 PA (CZAP/) CZAPLINSKI K.  
 PA (DINN/) DINMAN J D.  
 XX  
 XX Peltz S, Czaplinski K, Dinman JD;  
 PI WPI; 2004-449400/42.  
 DR  
 XX  
 XX Identifying a test composition or agent that modulates the efficiency of  
 PT translation termination comprises contacting the MTT1 with the test  
 PT composition or agent, and determining if the test composition or agent  
 PT inhibits the MTT1.  
 XX  
 XX Disclosure; SEQ ID NO 24; 41pp; English.  
 PS  
 XX  
 XX The invention relates to a method of identifying a test composition that  
 CC modulates the efficiency of translation termination comprising contacting  
 CC the RNA helicase MTT1 with a composition or agent under conditions  
 CC permitting binding between the MTT1 and the composition, detecting  
 CC specific binding of the test composition or agent to the MTT1, and  
 CC determining if the test composition or agent inhibits the MTT1. The  
 CC composition and methods are useful for modulating the fidelity of  
 CC translation termination or for identifying agents that: affect the  
 CC functional activity of mRNAs by altering frameshift frequency, permit  
 CC monitoring of a termination event, promote degradation of aberrant  
 CC transcripts, and provide modulators (inhibitors/stimulators) of peptidyl  
 CC transferase activity during initiation, elongation, termination and mRNA  
 CC degradation of translation. The agents, which may be antagonists or  
 CC agonists, are useful in screening, diagnostic and therapeutic purposes,  
 CC for diseases or conditions resulting from or cause premature translation,  
 CC such as beta-thalassemia, beta-globin, Duchenne/Becker Muscular  
 CC Dystrophy, Haemophilia A, Haemophilia B, Von Willebrand Disease,  
 CC Osteogenesis Imperfecta, Breast cancer, Ovarian Cancer, Wilms Tumour,  
 CC Hirschsprung disease, Cystic fibrosis, Kidney Stones, Familial  
 CC hypercholesterolaemia, Retinitis Pigmentosa, or Neurofibromatosis,  
 CC Retinoblastoma, ATM or Costmann Disease. The present sequence represents  
 CC the amino acid sequence of a yeast translation termination modulation  
 CC protein related peptide. Note: this sequence appears in the sequence  
 CC listing but no reference is made to it in the main body of the  
 CC specification.  
 XX  
 XX Sequence 11 AA;  
 SQ  
 Query Match 95.2%; Score 20; DB 8; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RRLN 4  
 ||||  
 DB  
 RESULT 44  
 ADA88834  
 ID ADA88834 standard; peptide; 12 AA.  
 XX  
 AC ADA88834;  
 XX  
 XX 20-NOV-2003 (first entry)  
 DT  
 XX Internalised peptide pep14 SEQ ID NO:14.  
 DE  
 XX internalising peptide; cytostatic; antiinflammatory; immunomodulator;  
 KW

Db 1 RRLN 4  
 RESULT 43  
 AAB59999  
 ID AAB59999 standard; peptide; 12 AA.  
 AC AAB59999;  
 XX  
 XX 05-NOV-2001 (first entry)  
 DT  
 XX Internalising peptide SEQ ID NO: 14.  
 DE  
 XX Internalising peptide; transport; apoptosis; arthritis; cancer;  
 KW stem cell; cell differentiation; immune response stimulation;  
 KW HIV vaccine.  
 XX  
 OS Synthetic.  
 XX  
 PN WO200115511-A2.  
 XX  
 XX 08-MAR-2001.  
 PD  
 XX  
 XX 31-AUG-2000; 2000WO-US024034.  
 PF  
 XX 01-SEP-1999; 99US-0151980P.  
 PR  
 XX 13-MAR-2000; 2000US-0188944P.  
 PR  
 XX (UYPI-) UNIV PITTSBURGH.  
 PA  
 XX  
 XX Robbins PD, Mi Z, Frizzell R, Glorioso JC, Gambotto A;  
 PI WPI; 2001-273309/28.  
 DR  
 XX Peptides that facilitate uptake and cytoplasmic and/or nuclear transport  
 PT of proteins, DNA and viruses, useful, e.g. for facilitating uptake of  
 PT antigens in immunogenic compositions.  
 PT  
 XX  
 XX Claim 1; Page 116; 129pp; English.  
 PS  
 XX The present invention provides the sequences of 75 peptides which  
 CC facilitate the uptake and transport of viruses, proteins and nucleic  
 CC acids. These internalising peptides can be used for transport into the  
 CC cytoplasm or the nucleus. They are useful for facilitating uptake into  
 CC the cell, inducing apoptosis, for example in the treatment of arthritis  
 CC and cancer, to expand a population of stem cells or differentiated cells,  
 CC to stimulate cell differentiation, facilitate the integration of AAV into  
 CC the genome of a cell, and to stimulate an immune response, for example in  
 CC the case of a HIV vaccine. The present sequence is one of the peptides of  
 CC the invention  
 XX  
 SQ Sequence 12 AA;  
 Query Match 95.2%; Score 20; DB 4; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RRLN 4  
 ||||  
 DB 8 RRLN 11  
 RESULT 44  
 ADA88834  
 ID ADA88834 standard; peptide; 12 AA.  
 XX  
 AC ADA88834;  
 XX  
 XX 20-NOV-2003 (first entry)  
 DT  
 XX Internalised peptide pep14 SEQ ID NO:14.  
 DE  
 XX internalising peptide; cytostatic; antiinflammatory; immunomodulator;  
 KW

KW antiarthritic; cytoplasmic transport; nuclear transport;  
 KW peptide-cargo complex; apoptosis; arthritic; tumour; differentiation;  
 KW immune response; vaccine; inflammation; necrosis; transplantation;  
 KW cystic fibrosis; lung inflammation; gene therapy.  
 XX  
 OS Synthetic.  
 XX  
 XX WO2003068942-A2.  
 XX  
 XX PD 21-AUG-2003.  
 XX  
 XX PF 12-FEB-2003; 2003WO-US004632.  
 XX  
 XX PR 13-FEB-2002; 2002US-00075869.  
 XX  
 XX PA (UYFI-) UNIV PITTSBURGH.  
 XX  
 XX PI Robbins PD, Mi Z, Frizzel R, Glorioso JC, Gambotto A, Mai JC;  
 XX  
 XX DR WPI; 2003-697526/66.  
 XX  
 XX PT New internalizing peptides, useful for facilitating the delivery, uptake  
 PT and cytoplasmic and/or nuclear transport of proteins, DNA or viruses into  
 PT a target cell, for inducing apoptosis in arthritic or tumor cells, or in  
 PT gene therapy.  
 XX  
 XX PS Example 3; Page 18; 171pp; English.  
 XX  
 CC The present invention describes an internalising peptide (I) comprising  
 CC any one of 14 fully defined amino acid sequences (designated PI-P14, see  
 CC ADA8896 to ADA8906, and ADA8917 to ADA8919). (I) has cytosolic,  
 CC antiinflammatory, immunomodulator and antiarthritic activities. The  
 CC internalising peptides are useful for facilitating the delivery, uptake  
 CC and cytoplasmic and/or nuclear transport of cargo, e.g. proteins, DNA or  
 CC viruses, into a target cell. The internalising peptides and peptide-cargo  
 CC complexes from the present invention are also useful for inducing  
 CC apoptosis in cells (e.g. arthritic cells or tumor cells), expanding a  
 CC population of stem cell or differentiated cells, stimulating the  
 CC differentiation of a population of stem cells, facilitating the  
 CC integration of adeno-associated virus DNA into the genome of a cell,  
 CC stimulating or eliciting an immune response in a subject, facilitating  
 CC the delivery of immunogens (e.g. vaccines), inhibiting the inflammatory  
 CC process, protecting tissue from apoptosis or necrosis during tissue  
 CC isolation prior to transplantation, facilitating transfer of proteins and  
 CC peptides to the lung for the treatment of cystic fibrosis or lung  
 CC inflammation, or in gene therapy. The present sequence represents a  
 CC peptide used in the exemplification of the present invention.  
 XX  
 XX SQ Sequence 12 AA;  
 Query Match 95.2%; Score 20; DB 7; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RRLN 4  
 Db 8 RRLN 11  
 RESULT 45  
 ADQ15921  
 ID ADQ15921 standard; peptide; 12 AA.  
 XX  
 XX AC ADQ15921;  
 XX  
 XX DT 07-OCT-2004 (first entry)  
 XX  
 XX DE Human Mash-1 splice variant hMash-1-deltaN76-specific epitope.  
 XX  
 KW Human; transcriptional modulator; splice variant; tumour-specific;  
 KW tumour-enriched; differential expression; expression pattern; cancer;  
 KW diagnosis; Mash-1; variant-specific antibody; autoantibody;  
 KW diagnostic array; prognosis; lung cancer; gastrointestinal cancer;

KW breast cancer; prostate cancer; skin cancer; sarcoma; endocrine cancer;  
 KW neural cancer; bladder cancer; cervical cancer; renal cancer;  
 KW haematopoietic cancer; cycostatic; transcription modulator inhibitor;  
 KW gene therapy; hMash-1-deltaN76; epitope.  
 XX  
 OS Homo sapiens.  
 XX  
 XX PN WO2004060302-A2.  
 XX  
 XX PD 22-JUL-2004.  
 XX  
 XX PF 24-DEC-2003; 2003WO-US041253.  
 XX  
 XX PR 26-DEC-2002; 2002US-0436693P.  
 XX  
 XX PA (CEMI-) CEMINES LLC.  
 XX  
 XX PI Neuman T, Palm K;  
 XX  
 XX DR WPI; 2004-534307/51.  
 XX  
 XX PT Transcription modulator splice variants, useful in diagnosing or treating  
 PT cancers such as lung, gastrointestinal, breast, prostate, skin,  
 PT endocrine, cervical and renal cancer.  
 XX  
 XX PS Claim 53; Page 50; 59pp; English.  
 XX  
 CC The invention relates to tumour-specific or tumour-enriched splice  
 CC variants of the human transcription modulators neuralised-1 (hNeu delta  
 CC NHR1), Neurodi (hNeurodi), Irx-2 (Irx-2A) and Mash-1 (hMash-1-deltaN76)  
 CC and nucleic acids encoding these splice variants. The invention also  
 CC relates to a method of diagnosing cancer by determining the expression of  
 CC at least one splice variant of each of a plurality of transcription  
 CC modulators, or a plurality of splice variants of at least one  
 CC transcription modulator, where the splice variants are differentially  
 CC expressed compared to the wild-type isoform of the corresponding  
 CC transcription modulator, and where the expression pattern of the splice  
 CC variants is indicative of cancer. In this method, the transcription  
 CC modulators whose splice variant expression is analysed are selected from  
 CC NR5F, MDM2 TSG, RREB, ZNF207, TTF-1, GTFIIIA, HES-6, HRY, MSX2, Neu,  
 CC NeuroD1, Mash-1 and Irx-2. The invention further relates to a method for  
 CC the treatment of cancer by administration of a bioactive agent capable of  
 CC inhibiting the activity of one or more of the transcriptional modulator  
 CC splice variants; an antibody which binds to a transcriptional modulator  
 CC splice variant but not to the wild-type isoform; a peptide which binds to  
 CC the splice variant-specific antibody; and a diagnostic array for the  
 CC diagnosis of cancer comprising peptides which are capable of binding to  
 CC splice variant-specific autoantibodies but not to wild-type isoform-  
 CC specific autoantibodies. The methods and compositions of the present  
 CC invention are useful for the diagnosis, prognosis and/or treatment of  
 CC cancer, such as lung cancer, gastrointestinal cancer, breast cancer,  
 CC prostate cancer, skin cancer, sarcoma, endocrine cancer, neural cancer,  
 CC bladder cancer, cervical cancer, renal cancer and haematopoietic cancer.  
 CC The present sequence represents a specifically claimed peptide which  
 CC binds to an antibody specific for the human tumour-specific/enriched Mash  
 CC -1 splice variant hMash-1-deltaN76.  
 XX  
 XX SQ Sequence 12 AA;  
 Query Match 95.2%; Score 20; DB 8; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RRLN 4  
 Db 5 RRLN 8  
 RESULT 46  
 AAR47006  
 ID AAR47006 standard; protein; 14 AA.  
 XX  
 XX AC AAR47006;

XX 25-MAR-2003 (revised)  
 DT 16-SEP-1994 (first entry)  
 XX  
 DE LAR protein position 1302-1316.  
 XX  
 KW Naturally-occurring; immunomodulatory protein; human; therapy; class I;  
 KW major histocompatibility complex; class II; allotype; type I diabetes;  
 KW autoimmune disease; rheumatoid arthritis; T-cell-mediated response;  
 KW multiple sclerosis; transplant rejection; vaccine; MHC.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO9404171-AL.  
 PN  
 XX  
 PD 03-MAR-1994.  
 XX  
 XX 11-AUG-1993; 93WO-US007545.  
 XX  
 PR 11-AUG-1992; 92US-00925460.  
 PR 15-JUN-1993; 93US-00077255.  
 XX  
 XX (HARD ) HARVARD COLLEGE.  
 PA  
 XX Urban RG, Chiciz RM, Vignali DA, Hedley ML, Stern LJ;  
 PI Strominger JL;  
 PI  
 XX WPI; 1994-082825/10.  
 DR  
 XX Novel immunomodulatory peptide(s) and nucleic acids - useful for  
 PT treatment of autoimmune diseases, transplant rejection and for  
 PT vaccination.  
 PT  
 XX Disclosure; Page 48; 139pp; English.  
 PS  
 XX The sequences given in AAR49291-505 and AAR46981-7038 represent peptide  
 CC fragments of naturally-occurring immunomodulatory proteins. These  
 CC fragments are between 10-30 residues in length and bind to a human major  
 CC histocompatibility complex (MHC) class II allotype. These peptides may be  
 CC used for therapy of autoimmune diseases, such as type I diabetes,  
 CC rheumatoid arthritis and multiple sclerosis, and to reduce transplant  
 CC rejection. They may also be used for vaccination providing an exclusively  
 CC T-cell-mediated response, which can be class I or class-II based, or  
 CC both, depending on the length and character of the immunogenic peptides.  
 CC (Updated on 25-MAR-2003 to correct PN field.) (Updated on 25-MAR-2003 to  
 CC correct PR field.)  
 CC  
 XX Sequence 14 AA;  
 SQ  
 Query Match 95.2%; Score 20; DB 2; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RRLN 4  
 DB 6 RRLN 9  
 RESULT 47  
 AAW42267  
 ID AAW42267 standard; peptide; 15 AA.  
 XX  
 AC AAW42267;  
 XX  
 XX 08-APR-1998 (first entry)  
 DT  
 XX Biotinylated cross-linked interleukin-8 15-mer peptide ligand 7.  
 DE  
 XX Bacteriophage peptide library; peptide epitope; therapeutic target;  
 KW variegated compound library; interleukin-8; IL-8.  
 KW  
 OS Synthetic.  
 XX

PN WO9735194-A2.  
 XX  
 PD 25-SEP-1997.  
 XX  
 XX 21-MAR-1997; 97WO-US004176.  
 PR  
 XX 21-MAR-1996; 96US-00622338.  
 XX  
 XX (HARD ) HARVARD COLLEGE.  
 PA  
 XX Forster AC;  
 PI  
 XX WPI; 1997-480355/44.  
 DR  
 XX Identifying compounds which interact with target molecules - using  
 PT enantiomers of the target molecules and testing of enantiomers of  
 PT selected compounds.  
 PT  
 XX Disclosure; Fig 6; 89pp; English.  
 PS  
 XX 15-mer peptides AAW42261-77 are identified as ligands of a biotinylated,  
 CC cross-linked interleukin-8 (IL-8) target, using the method of the  
 CC invention. This novel method identifies compounds which interact with a  
 CC target molecule, and comprises contacting a screening molecule with a  
 CC variegated compound library, where the screening molecule comprises solid  
 CC target molecule, or the enantiomer if the target molecule is chiral.  
 CC Compounds which have a desired interaction with the target molecule are  
 CC selected, and the ability of their enantiomer to interact with the target  
 CC molecule is tested. Ligands for a target protein can be identified by  
 CC combining a D-enantiomer of a target protein (a D-target protein), and a  
 CC variegated compound library, and then selecting one or more compounds  
 CC from the library which have a desired binding interaction with the D-  
 CC target protein. The methods can be used for identifying agonists or  
 CC antagonists of targets such as receptors, enzymes, DNA binding proteins  
 CC or signal transduction proteins. The methods can provide a structurally  
 CC selective approach in addition to scoring for interaction of functional  
 CC groups. They provide a powerful selection method that allows for the  
 CC production of ligands with the same diversity as peptides but with the  
 CC greatly improved pharmacokinetic profiles needed for drug activity  
 CC  
 XX Sequence 15 AA;  
 SQ  
 Query Match 95.2%; Score 20; DB 2; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 4e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RRLN 4  
 DB 5 RRLN 8  
 RESULT 48  
 AAW15509  
 ID AAW15509 standard; peptide; 15 AA.  
 XX  
 AC AAW15509;  
 XX  
 XX 04-JUN-1997 (first entry)  
 DT  
 XX Peptide MIPP4.  
 DE  
 XX Phosphorylated PS peptide; antibody; phosphoenzyme.  
 KW  
 XX Synthetic.  
 OS  
 XX JP09068527-A.  
 PN  
 XX 11-MAR-1997.  
 PD  
 XX 31-AUG-1995; 95JP-00224179.  
 XX  
 XX 31-AUG-1995; 95JP-00224179.  
 XX

PA (IGAK-) IGAKU SEIBUTSUGAKU KENKYUSHO KK.  
 XX WPI; 1997-222835/20.  
 XX Monoclonal antibody against phosphorylated PS peptide - used in an  
 PT immunoassay for determining protein phospho:enzyme activity.  
 XX  
 XX Example 4; Page 10; 13pp; Japanese.  
 PS  
 XX This sequence represents the MPP4 peptide. This sequence is used in an  
 CC immunoassay method for determining protein phospho:enzyme activity. The  
 CC immunoassay method also uses a monoclonal antibody specific for a  
 CC phosphorylated PS peptide (see AAW15508). This method can determine the  
 CC protein phospho:enzyme activity in a sample with high specificity  
 XX  
 XX Sequence 15 AA;  
 SQ  
 Query Match 95.2%; Score 20; DB 2; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 4e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RRLN 4  
 DB 2 RRLN 5  
 RESULT 49  
 ABB76940  
 ID ABB76940 standard; peptide; 15 AA.  
 XX  
 AC ABB76940;  
 DT 22-JUL-2002 (first entry)  
 XX  
 DE Rat VG51-2 peptide.  
 XX  
 KW Rat; antiasthmatic; anxiolytic; antiepileptic; antihypertensive;  
 KW psychotropic; glutamate transporter; transporter; GABA;  
 KW gamma-aminobutyric acid transporter; GABA transporter; neurotransmitter;  
 KW asthma; anxiety; epilepsy; hypertension; psychiatric disorder;  
 KW neurotic disorder; VG51.  
 XX  
 OS Rattus sp.  
 XX  
 PN WO200071709-A1.  
 XX  
 PD 30-NOV-2000.  
 XX  
 PF 19-MAY-2000; 2000WO-FR001383.  
 XX  
 PR 21-MAY-1999; 99PR-00006525.  
 XX  
 PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.  
 XX  
 PI Giros B, Gasnier B, Sagne C, El Mestikawy S, Hamon M;  
 XX  
 DR WPI; 2001-025160/03.  
 XX  
 XX New mammalian amino acid transporter, used e.g. to screen for  
 PT psychotropic agents, is high capacity but low affinity transporter of  
 PT gamma-aminobutyric acid.  
 XX  
 PS Claim 3; Page 44; 103pp; French.  
 XX  
 CC The present sequence is a peptide of rat VG51, a glutamate/ gamma-  
 CC aminobutyric acid (GABA) transporter. GABA and glutamate are  
 CC neurotransmitters. The transporter can be used to produce specific  
 CC antibodies, to screen for binding agents. Modulators of the transporter  
 CC are useful for treating disorders associated with deregulated  
 CC glutamate/GABA transport, e.g. asthma, anxiety, epilepsy, hypertension  
 CC and other psychiatric and neurotic disorders, while determining levels of  
 CC the transporter and its coding sequence can be used for diagnosis of such  
 CC disorders

---

XX SQ Sequence 15 AA;  
 Query Match 95.2%; Score 20; DB 4; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 4e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RRLN 4  
 DB 1 RRLN 4  
 RESULT 50  
 AAE18774  
 ID AAE18774 standard; peptide; 15 AA.  
 XX  
 AC AAE18774;  
 XX  
 DT 17-MAY-2002 (first entry)  
 XX  
 DE Human leucocyte antigen (HLA) class II epitope #15.  
 XX  
 KW Human; lung tumour associated antigen; CASB761; vaccine; lung cancer;  
 KW immunotherapeutic; lung preneoplastic lesion; autoimmune disease;  
 KW Gene therapy; cycostatic; immunosuppressive; human leucocyte antigen;  
 KW HLA; epitope.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200206338-A1.  
 XX  
 PD 24-JAN-2002.  
 XX  
 PF 11-JUL-2001; 2001WO-EP007967.  
 XX  
 PR 17-JUL-2000; 2000GB-00017512.  
 XX  
 PA (SMTK) SMITHKLINE BEECHAM BIOLOGICALS.  
 XX  
 PI Cassart J, Gaulis S, Vinals Y De Bassols C;  
 XX  
 DR WPI; 2002-179782/23.  
 XX  
 XX Vaccine composition for treating cancer, in particular lung cancer, tumor  
 PT autoimmune diseases and other related conditions, comprises a lung tumor  
 PT associated antigen, especially CASB761 polypeptide.  
 XX  
 PS Claim 9; Page 67; 92pp; English.  
 XX  
 CC The invention relates to vaccines comprising lung tumour associated  
 CC antigen referred as CASB761 and its polynucleotide. CASB761 and its DNA  
 CC are useful in the manufacture of a vaccine for immunotherapeutically  
 CC treating a patient suffering from or susceptible to lung cancer, lung  
 CC preneoplastic lesions or other related conditions. Vaccines of the  
 CC invention are useful in medicine, for treating cancer, particularly lung  
 CC cancer, autoimmune diseases and other related conditions. CASB761  
 CC polynucleotides and their proteins are useful as diagnostic reagents, to  
 CC diagnose and grading the nature of the cancerous tissue. An antibody  
 CC expressing CASB761 protein or to treat or prevent, particularly lung cancer,  
 CC immunospecific for CASB761 is useful to isolate and to identify clones  
 CC chromatography and to treat or prevent, particularly lung cancer,  
 CC autoimmune disease and related conditions. CASB761 DNA is used in gene  
 CC therapy. The present sequence is human leucocyte antigen (HLA) class II  
 CC epitope. This sequence is used to incorporate an epitope of CASB761  
 CC protein  
 XX SQ Sequence 15 AA;  
 Query Match 95.2%; Score 20; DB 5; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 4e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
Db ||||  
8 RRLN 11

Search completed: January 25, 2006, 18:39:30  
Job time : 129.5 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:34:13 ; Search time 12.5 Seconds  
(without alignments)  
38.487 Million cell updates/sec

Title: US-10-771-242-295

Perfect score: 21

Sequence: 1 RRLNX 5

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database : PIR 80:\*

1: pirl:\*

2: pirl:\*

3: pirl:\*

4: pirl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	95.2	51	2 AF0749	hypothetical prote
2	20	95.2	51	2 S00016	protamine Z1 - sma
3	20	95.2	51	2 D98179	hypothetical prote
4	20	95.2	61	2 B84091	hypothetical prote
5	20	95.2	71	2 B64479	hypothetical prote
6	20	95.2	72	2 I39470	apolipoprotein B -
7	20	95.2	73	2 E90729	hypothetical prote
8	20	95.2	73	2 B90775	C4-type zinc finger
9	20	95.2	73	2 F91004	probable C4-type z
10	20	95.2	74	2 E82829	hypothetical prote
11	20	95.2	76	2 C64871	hypothetical prote
12	20	95.2	80	2 C91287	hypothetical prote
13	20	95.2	85	2 T08595	cysteine proteinas
14	20	95.2	87	2 S00180	spermatid protein
15	20	95.2	87	2 B37475	probable structura
16	20	95.2	89	2 C90392	conserved hypotet
17	20	95.2	90	2 H83003	conserved hypotet
18	20	95.2	91	2 A95415	hypothetical prote
19	20	95.2	92	2 AE0430	hypothetical prote
20	20	95.2	94	1 RBP22	abcl protein A pha
21	20	95.2	95	2 AD0229	hypothetical prote
22	20	95.2	96	2 B26074	cysteine proteinas
23	20	95.2	96	2 S51929	homeotic protein C
24	20	95.2	96	2 S51928	homeotic protein C
25	20	95.2	98	2 S62346	L71-5 protein - fr
26	20	95.2	98	2 A39437	exopolysaccharide
27	20	95.2	98	2 F95975	posttranscription
28	20	95.2	98	2 G96029	hypothetical prote
29	20	95.2	100	2 S62333	L71-1 protein - fr

30	20	95.2	102	2 T50981	hypothetical prote
31	20	95.2	103	2 E82425	hypothetical prote
32	20	95.2	103	2 S43813	hypothetical prote
33	20	95.2	105	2 T06820	lipid transfer pro
34	20	95.2	108	2 F84479	En/Spn-like transp
35	20	95.2	109	1 CABO	cytochrome-c oxida
36	20	95.2	110	2 AF0301	conserved hypotet
37	20	95.2	112	2 F71149	hypothetical prote
38	20	95.2	113	2 E83073	hypothetical prote
39	20	95.2	113	2 AD2868	hypothetical prote
40	20	95.2	114	2 AI0294	probable membrane
41	20	95.2	115	2 AF0052	hypothetical prote
42	20	95.2	115	2 S64490	hypothetical prote
43	20	95.2	116	2 T18073	hypothetical prote
44	20	95.2	118	2 H97644	hypothetical prote
45	20	95.2	121	2 B72663	hypothetical prote
46	20	95.2	121	2 G70074	hypothetical prote
47	20	95.2	123	2 G01477	ribosomal protein
48	20	95.2	123	2 B95328	hypothetical prote
49	20	95.2	124	2 T37118	probable transposa
50	20	95.2	125	2 AF2757	conserved hypotet
51	20	95.2	125	2 C97538	hypothetical prote
52	20	95.2	126	2 S47103	urease (EC 3.5.1.5
53	20	95.2	128	2 A72533	hypothetical prote
54	20	95.2	128	2 H65171	hypothetical trans
55	20	95.2	128	2 C84708	probable MYB fami
56	20	95.2	129	2 A43952	neuropeptides prec
57	20	95.2	129	2 B71068	hypothetical prote
58	20	95.2	130	2 A83400	hypothetical prote
59	20	95.2	131	2 AF0510	transcription acti
60	20	95.2	134	2 S19144	spba protein - Bac
61	20	95.2	134	2 S77785	hypothetical prote
62	20	95.2	135	2 S75912	hypothetical prote
63	20	95.2	136	2 H71668	ribosomal protein
64	20	95.2	138	2 F81898	probable membrane
65	20	95.2	139	1 RKMLS	ribulose-bisphosph
66	20	95.2	139	2 D90303	hypothetical prote
67	20	95.2	140	2 T09267	probable tail comp
68	20	95.2	140	2 G85746	unknown protein en
69	20	95.2	140	2 T09072	hypothetical prote
70	20	95.2	141	2 T09198	probable tail comp
71	20	95.2	144	2 B83596	conserved hypotet
72	20	95.2	144	2 F28585	protein D - Escher
73	20	95.2	145	1 S05495	cytochrome-c oxida
74	20	95.2	145	1 R3HU15	ribosomal protein
75	20	95.2	145	1 R3HY15	ribosomal protein
76	20	95.2	145	1 R3RM15	ribosomal protein
77	20	95.2	145	2 B34823	ribosomal protein
78	20	95.2	145	2 C34823	ribosomal protein
79	20	95.2	145	2 A34823	ribosomal protein
80	20	95.2	146	1 S04592	cytochrome-c oxida
81	20	95.2	146	2 A95254	adc operon repress
82	20	95.2	146	2 E84608	hypothetical prote
83	20	95.2	146	2 T46753	repressor protein
84	20	95.2	147	2 G98091	conserved hypotet
85	20	95.2	147	2 C95227	conserved hypotet
86	20	95.2	147	2 T49502	hypothetical prote
87	20	95.2	148	2 S48980	hypothetical prote
88	20	95.2	149	2 F97312	hypothetical prote
89	20	95.2	150	1 OTHU5A	cytochrome-c oxida
90	20	95.2	150	2 F81280	hypothetical prote
91	20	95.2	151	2 G70377	conserved hypotet
92	20	95.2	151	2 A81094	conserved hypotet
93	20	95.2	151	2 B81847	conserved hypotet
94	20	95.2	152	2 A70800	hypothetical prote
95	20	95.2	156	2 T17274	hypothetical prote
96	20	95.2	160	2 A71300	conserved hypotet
97	20	95.2	161	2 S30937	allophycocyanin al
98	20	95.2	161	2 A30938	myosin regulatory
99	20	95.2	161	2 E71866	hypothetical prote
100	20	95.2	162	2 E83118	probable phenazine
101	20	95.2	163	2 T22481	hypothetical prote
102	20	95.2	164	2 A82393	hypothetical prote

103	20	95.2	165	2	AB1760	acetyltransferase	176	20	95.2	214	2	C42327	alpha 2,6-sialyltr
104	20	95.2	166	2	G98118	hypothetical prote	177	20	95.2	217	2	S65830	alpha fucosidase p
105	20	95.2	167	2	AB2180	DnaJ protein [limpo	178	20	95.2	217	2	S49578	crystin inhibitor
106	20	95.2	168	2	I40352	single-stranded DN	179	20	95.2	217	2	JQ1294	hypothetical 23.7k
107	20	95.2	168	2	AB3362	single-strand bind	180	20	95.2	218	2	T27954	hypothetical prote
108	20	95.2	168	2	B83473	hypothetical prote	181	20	95.2	218	2	S64086	hypothetical prote
109	20	95.2	169	2	AB1032	hypothetical prote	182	20	95.2	219	2	A96813	hypothetical prote
110	20	95.2	170	2	D71141	hypothetical prote	183	20	95.2	219	2	I51382	achaete-scute homo
111	20	95.2	170	2	T25399	hypothetical prote	184	20	95.2	220	2	S31152	hypothetical prote
112	20	95.2	170	2	T34588	hypothetical prote	185	20	95.2	220	2	D64316	restriction modifi
113	20	95.2	171	2	H75010	hypothetical prote	186	20	95.2	221	2	T33414	hypothetical prote
114	20	95.2	171	2	H87618	hypothetical prote	187	20	95.2	222	2	T37842	hypothetical prote
115	20	95.2	172	2	AD2231	dihydrofolate redu	188	20	95.2	222	2	T20534	conserved hypotet
116	20	95.2	174	2	AD2665	conserved hypotet	189	20	95.2	224	2	E72049	CT691 hypothet
117	20	95.2	175	2	T31918	hypothetical prote	190	20	95.2	224	2	G86575	hypothetical prote
118	20	95.2	176	2	F71064	micrococcal nuclea	191	20	95.2	224	2	G90448	hypothetical prote
119	20	95.2	176	2	B72698	hypothetical prote	192	20	95.2	225	2	B97580	hypothetical prote
120	20	95.2	177	2	G64499	hypothetical prote	193	20	95.2	225	2	AH2800	conserved hypotet
121	20	95.2	177	2	T36271	probable RNA polym	194	20	95.2	225	2	C70500	probable transcrip
122	20	95.2	179	2	G76208	CPDdiacylglycerol-	195	20	95.2	226	2	JC5327	adhesin complex 25
123	20	95.2	179	2	A12104	filament integrity	196	20	95.2	227	2	AD3529	two component resp
124	20	95.2	181	2	S77590	cold shock protein	197	20	95.2	227	2	T32894	hypothetical prote
125	20	95.2	184	2	S41451	regA protein - Rho	198	20	95.2	228	2	S76711	hypothetical prote
126	20	95.2	184	2	A42219	light-harvesting a	199	20	95.2	228	2	T51147	hypothetical prote
127	20	95.2	184	2	A36862	photosynthetic res	200	20	95.2	228	2	AF3179	agrobacterium viru
128	20	95.2	185	2	D83658	stage V sporulatio	201	20	95.2	229	1	G64371	conserved hypotet
129	20	95.2	185	2	A70075	hypothetical prote	202	20	95.2	229	2	H82043	guanylate kinase V
130	20	95.2	185	2	G70425	hypothetical prote	203	20	95.2	229	2	D90668	probable xanthine
131	20	95.2	185	2	F87279	response regulator	204	20	95.2	229	2	F64754	probable oxidoredu
132	20	95.2	187	2	S32968	probable membrane	205	20	95.2	229	2	G85518	glucuronolactone r
133	20	95.2	187	2	AB3478	acid tolerance reg	206	20	95.2	229	2	T05150	hypothetical prote
134	20	95.2	189	2	G70485	hypothetical prote	207	20	95.2	230	2	C75412	probable cytochrom
135	20	95.2	189	2	B69177	hypothetical prote	208	20	95.2	230	2	S28186	achaete-scute locu
136	20	95.2	191	1	VU0B1B	BabR protein 1 - B	209	20	95.2	231	2	T51160	hypothetical prote
137	20	95.2	192	2	S49429	hypothetical prote	210	20	95.2	233	2	S11563	probable MASH-2 pr
138	20	95.2	195	2	B81066	hypothetical prote	211	20	95.2	233	2	H95253	zinc ABC transport
139	20	95.2	195	2	AC2583	two component resp	212	20	95.2	234	2	F98118	hypothetical prote
140	20	95.2	195	2	A97365	actR protein (AF22	213	20	95.2	234	2	T46754	AdcC protein limpo
141	20	95.2	196	2	G75353	hypothetical prote	214	20	95.2	235	2	S70132	probable membrane
142	20	95.2	196	2	F70428	probable CDP-alcoh	215	20	95.2	235	2	AG2170	hypothetical prote
143	20	95.2	196	2	I50507	achaete-scute homo	216	20	95.2	237	2	G97418	hypothetical prote
144	20	95.2	197	1	YTBS7	tunicamycin resist	217	20	95.2	237	2	AC2913	transcription regu
145	20	95.2	198	2	C97447	hypothetical prote	218	20	95.2	238	2	F97687	hypothetical prote
146	20	95.2	198	2	T25436	hypothetical prote	219	20	95.2	238	2	D86917	conserved hypotet
147	20	95.2	199	2	A56548	pro-neural achaete	220	20	95.2	238	2	A48279	achaete scute prot
148	20	95.2	199	2	VU0B2B	BabR protein 2 - B	221	20	95.2	239	2	S07406	thaumatin homolog
149	20	95.2	200	1	B83558	hypothetical prote	222	20	95.2	239	2	B87280	conserved hypotet
150	20	95.2	200	2	AB2123	hypothetical prote	223	20	95.2	239	2	AF0555	probable regulator
151	20	95.2	201	2	A97327	probable membrane	224	20	95.2	239	2	T46948	probable repressor
152	20	95.2	202	2	S39198	nitrogenase (BC 1.	225	20	95.2	239	2	E83983	tetracyclodipicoli
153	20	95.2	202	2	S39197	nitrogenase (BC 1.	226	20	95.2	240	2	A95219	hypothetical prote
154	20	95.2	202	2	C83859	hypothetical prote	227	20	95.2	240	2	G98082	hypothetical prote
155	20	95.2	203	2	F90302	conserved hypotet	228	20	95.2	240	2	D90337	hypothetical prote
156	20	95.2	204	2	AC1361	ATP-dependent Clp	229	20	95.2	240	2	E84701	hypothetical prote
157	20	95.2	205	2	F81718	guanylate kinase T	230	20	95.2	241	2	AC3083	phosphonate ABC tr
158	20	95.2	205	2	A71567	probable GMP kinas	231	20	95.2	241	2	S75347	hypothetical prote
159	20	95.2	205	2	F86932	conserved hypotet	232	20	95.2	241	2	T08536	transfer origin pr
160	20	95.2	205	2	C88288	protein ZK970.2 [i	233	20	95.2	241	2	S23004	trah protein - Esc
161	20	95.2	206	2	A83318	hypothetical prote	234	20	95.2	241	2	S59377	probable membrane
162	20	95.2	207	2	AD3229	protein 6b [import	235	20	95.2	242	2	A86719	oxidoreductase yhg
163	20	95.2	207	2	B30832	hypothetical prote	236	20	95.2	242	2	A95180	conserved hypotet
164	20	95.2	208	2	S59772	hypothetical prote	237	20	95.2	242	2	E89813	hypothetical prote
165	20	95.2	209	2	G96969	probable methyltra	238	20	95.2	242	2	D98047	conserved hypotet
166	20	95.2	209	2	C83079	hypothetical prote	239	20	95.2	243	2	AB7478	methyltransferase,
167	20	95.2	210	2	F87492	ATP-dependent Clp	240	20	95.2	243	2	C75605	hypothetical prote
168	20	95.2	210	2	E97512	clpp (AF218420) [i	241	20	95.2	244	2	S50685	helicase-related p
169	20	95.2	210	2	AE2731	ATP-dependent Clp	242	20	95.2	244	2	D72258	ABC-type multidrug
170	20	95.2	210	2	A95855	probable transcrip	243	20	95.2	245	2	F97002	endopeptidase Clp
171	20	95.2	210	2	A87022	conserved hypotet	244	20	95.2	245	2	AD3361	osmotin-like prote
172	20	95.2	212	1	VU0B3B	BabR protein 3 - B	245	20	95.2	247	2	S33196	osmotin-like prote
173	20	95.2	212	2	T26887	hypothetical prote	246	20	95.2	247	2	S33197	osmotin-like prote
174	20	95.2	213	2	T27841	hypothetical prote	247	20	95.2	248	2	S60940	RSC complex chain
175	20	95.2	213	2	AH2636	conserved hypotet	248	20	95.2	248	2		

249	20	95.2	248	2	T44922	hypothetical prote	322	20	95.2	282	2	AG3545	dipeptide transpor
250	20	95.2	249	2	S04674	hypothetical prote	323	20	95.2	282	2	T18608	hypothetical prote
251	20	95.2	249	2	F91274	hypothetical prote	324	20	95.2	282	2	E84136	ABC transporter (A
252	20	95.2	249	2	F86115	hypothetical prote	325	20	95.2	282	2	F86396	hypothetical prote
253	20	95.2	249	2	S56415	hypothetical 27.6K	326	20	95.2	283	2	AD2602	conserved hypotet
254	20	95.2	250	1	OBNC2	cytochrome-c oxida	327	20	95.2	283	2	D97384	endopeptidase-rela
255	20	95.2	250	1	C45353	coat protein - app	328	20	95.2	284	1	E64938	probable aldehyde
256	20	95.2	250	2	D71061	hypothetical prote	329	20	95.2	284	2	F85788	probable an aldehy
257	20	95.2	252	1	WMBV2P	29K protein - toba	330	20	95.2	284	2	B90940	probable an aldehy
258	20	95.2	252	2	D81052	hypothetical prote	331	20	95.2	284	2	D81075	conserved hypotet
259	20	95.2	252	2	AH3267	acyltransferase, p	332	20	95.2	284	2	D81866	hypothetical prote
260	20	95.2	252	2	S61142	hypothetical prote	333	20	95.2	284	2	G82319	DNAJ-related prote
261	20	95.2	252	2	B97582	phosphoribosylamin	334	20	95.2	285	2	E85521	polysaccharide hyd
262	20	95.2	254	2	A12802	hypothetical prote	335	20	95.2	285	2	A71101	transport protein
263	20	95.2	254	2	AD3392	phosphoribosylamin	336	20	95.2	285	2	AG1535	hypothetical prote
264	20	95.2	254	2	F90408	conserved hypotet	337	20	95.2	286	2	B83259	hypothetical prote
265	20	95.2	255	2	I48235	inhibin beta-B cha	338	20	95.2	287	1	S49944	actin-capping prot
266	20	95.2	255	2	F97235	probable PHP famil	339	20	95.2	287	2	S26190	nitrogenase (EC 1.
267	20	95.2	255	2	JQ3232	coat protein - ind	340	20	95.2	287	2	B95016	transcription regu
268	20	95.2	256	2	S22589	hypothetical prote	341	20	95.2	287	2	B86324	protein Fl4D16.21
269	20	95.2	256	2	D96743	hypothetical prote	342	20	95.2	287	2	D97889	positive transcrip
270	20	95.2	256	2	D96743	unknown protein [i	343	20	95.2	288	2	T45536	moar protein [impo
271	20	95.2	257	2	JQ3232	AR1 protein - cas	344	20	95.2	288	2	S16556	hypothetical prote
272	20	95.2	257	2	S39234	gene V2 protein -	345	20	95.2	289	2	AH3495	gamma-D-glutamyl-L
273	20	95.2	257	2	T40880	hypothetical prote	346	20	95.2	291	2	C81943	probable integrase
274	20	95.2	257	2	AH0835	probable reverse t	347	20	95.2	291	2	D89839	hypothetical prote
275	20	95.2	258	1	QQMC1	coat protein - cas	348	20	95.2	291	2	AC0085	probable membrane
276	20	95.2	258	1	VCMCN	coat protein - cas	349	20	95.2	292	2	S11870	peroxidase (EC 1.1
277	20	95.2	258	2	S25624	coat protein - cas	350	20	95.2	293	2	AE0372	dihydrodipicolinat
278	20	95.2	258	2	S25622	coat protein - cas	351	20	95.2	293	2	S72970	tyrosine-tRNA liga
279	20	95.2	258	2	S25623	coat protein - cas	352	20	95.2	293	2	AI0404	probable hexulose-
280	20	95.2	258	2	T02819	probable membrane	353	20	95.2	294	2	G82483	probable hexulose-
281	20	95.2	258	2	G82868	hypothetical prote	354	20	95.2	294	2	H96719	homeobox gene 13 p
282	20	95.2	259	2	F69678	involved in polyke	355	20	95.2	295	1	P3WMBB	3a protein - broad
283	20	95.2	259	2	T13260	hypothetical prote	356	20	95.2	295	2	E96736	probable dehydroge
284	20	95.2	259	2	D86685	prophage p11 prote	357	20	95.2	297	2	T36724	probable membrane
285	20	95.2	259	2	C86797	prophage p13 prote	358	20	95.2	297	2	AE1816	ABC transporter (A
286	20	95.2	259	2	C86757	prophage p12 prote	359	20	95.2	298	2	B75612	phosphate ABC tran
287	20	95.2	260	1	AC30202	transcription init	360	20	95.2	299	2	D96736	probable dehydroge
288	20	95.2	260	1	QQCVCL	coat protein - tom	361	20	95.2	299	2	C87347	transcription regu
289	20	95.2	260	2	A64561	NH(3)-dependent NA	362	20	95.2	299	2	AD3083	transcription regu
290	20	95.2	260	2	AE0438	hypothetical prote	363	20	95.2	300	2	F70196	flagellar hook-bas
291	20	95.2	261	2	C86481	30.5K hypothetical	364	20	95.2	301	2	B64189	glycine cleavage s
292	20	95.2	262	2	D85753	probable transient	365	20	95.2	301	2	E90608	ABC transporter at
293	20	95.2	262	2	E90865	probable transient	366	20	95.2	301	2	T14331	homeotic protein -
294	20	95.2	264	2	A11802	an E. coli prôtein	367	20	95.2	301	2	B86966	probable oxidoredu
295	20	95.2	264	2	AB1429	an B. coli prôtein	368	20	95.2	302	2	A47126	alsSD operon activ
296	20	95.2	264	2	F72714	hypothetical prote	369	20	95.2	302	2	B49941	devR protein - Myx
297	20	95.2	265	2	E64880	hypothetical prote	370	20	95.2	303	2	D85356	cinnamoyl-CoA redu
298	20	95.2	266	2	G90785	probable acetyltra	371	20	95.2	304	2	G95274	probable gluconola
299	20	95.2	266	2	E85645	probable acetyltra	372	20	95.2	304	2	A95364	protein [imported
300	20	95.2	266	2	G64842	probable hydrolase	373	20	95.2	305	2	B90136	hypothetical prote
301	20	95.2	267	1	A46535	interleukin-2, rece	374	20	95.2	305	2	C87356	conserved hypotet
302	20	95.2	267	2	S76499	hypothetical prote	375	20	95.2	306	2	C87283	glutaminase A limp
303	20	95.2	267	2	G90467	conserved hypotet	376	20	95.2	307	2	B81874	hypothetical prote
304	20	95.2	268	2	T04966	hypothetical prote	377	20	95.2	307	2	C81148	YabO/YceC/SfhB fam
305	20	95.2	268	2	T38394	probable methyl tr	378	20	95.2	308	2	C95406	hypothetical prote
306	20	95.2	270	2	F64460	hypothetical prote	379	20	95.2	309	1	EUSMAG	agatase (EC 3.2.1.
307	20	95.2	271	2	H83622	hypothetical prote	380	20	95.2	309	2	F82936	thioredoxin reduct
308	20	95.2	272	2	I48700	gene ox40 protein	381	20	95.2	309	2	A83883	sugar transport sy
309	20	95.2	272	2	B75293	amino acid ABC tra	382	20	95.2	310	2	C72529	hypothetical prote
310	20	95.2	273	2	C72328	transaminase B hom	383	20	95.2	310	2	F96928	ABC-type multidrug
311	20	95.2	273	2	AD1855	ATP-binding protei	384	20	95.2	310	2	C96929	ABC-type multidrug
312	20	95.2	273	2	G81952	HemK protein NMA03	385	20	95.2	311	1	LNHU2A	asialoglycoprotein
313	20	95.2	273	2	AF3017	flagellar motor pr	386	20	95.2	311	2	H95877	hypothetical prote
314	20	95.2	276	2	JCS285	carbonyl reductase	387	20	95.2	311	2	B72781	hypothetical prote
315	20	95.2	276	2	JQ2340	anthranilate synth	388	20	95.2	312	2	H86732	hypothetical prote
316	20	95.2	276	2	F84859	hypothetical prote	389	20	95.2	312	2	AB0306	probable membrane
317	20	95.2	276	2	AB2037	hypothetical prote	390	20	95.2	312	2	H97342	ABC-type MDR trans
318	20	95.2	278	2	G64393	hypothetical prote	391	20	95.2	312	2	G97260	ABC-type MDR trans
319	20	95.2	279	2	AB0285	probable regulator	392	20	95.2	312	2	S49612	transposase - Esch
320	20	95.2	280	2	AI1856	hypothetical prote	393	20	95.2	313	2	E83616	probable transcrip
321	20	95.2	281	2	T22670	hypothetical prote	394	20	95.2	313	2	G97006	ABC-type multidrug

395	20	95.2	313	2	T12634	homeotic protein -	468	20	95.2	347	1	A44245	alcohol dehydrogen
396	20	95.2	314	2	T32380	hypothetical prote	469	20	95.2	347	1	S51120	hypothetical prote
397	20	95.2	314	2	D97318	ABC-type MDR trans	470	20	95.2	347	2	AG2433	hypothetical prote
398	20	95.2	315	2	D96948	2-keto-3-deoxygluc	471	20	95.2	348	2	A82322	octaprenyl-diphosp
399	20	95.2	316	2	D75588	transcriptional regu	472	20	95.2	349	1	WFFGBB	inhibin beta-B cha
400	20	95.2	316	2	B71082	hypothetical prote	473	20	95.2	350	2	G95382	probable Transport
401	20	95.2	317	2	F83730	hypothetical prote	474	20	95.2	351	2	AE2979	aldo/keto reductas
402	20	95.2	318	2	H81255	hypothetical prote	475	20	95.2	351	2	H98303	hypothetical prote
403	20	95.2	318	2	S52424	homeodomain protei	476	20	95.2	353	1	H71340	membrane lipoprote
404	20	95.2	319	2	AG0114	glutathione syntha	477	20	95.2	354	2	T46875	Dbh protein (impor
405	20	95.2	319	2	T34728	hypothetical prote	478	20	95.2	356	2	A91023	isochorismate hydr
406	20	95.2	320	2	A84651	hypothetical prote	479	20	95.2	356	2	G64997	isochorismate synt
407	20	95.2	321	2	T13539	hypothetical prote	480	20	95.2	356	2	B85867	hypothetical prote
408	20	95.2	322	2	S62736	cathepsin-like cys	481	20	95.2	356	2	T41764	AcMNPV orf18 - Bom
409	20	95.2	323	2	JCS691	cysteine proteinas	482	20	95.2	357	2	T17471	hypothetical prote
410	20	95.2	323	2	T22956	hypothetical prote	483	20	95.2	357	2	D83385	probable (S)-2-hyd
411	20	95.2	323	2	S59373	cyclin homolog UME	484	20	95.2	358	2	S41640	ribosomal protein
412	20	95.2	324	2	S62735	cathepsin - Choriz	485	20	95.2	359	2	D83385	D-vitopine dehydro
413	20	95.2	325	2	E89939	Heat-inducible tra	486	20	95.2	360	1	S30109	DNA topoisomerase
414	20	95.2	326	2	C91228	hypothetical prote	487	20	95.2	360	2	D59367	hypothetical prote
415	20	95.2	326	2	B86075	hypothetical prote	488	20	95.2	360	2	A12042	cysteine proteinas
416	20	95.2	326	2	S40818	hypothetical 36.9K	489	20	95.2	361	2	B84601	probable N-acetylgl
417	20	95.2	329	2	T18617	hypothetical prote	490	20	95.2	361	2	F97068	histidinol-phospha
418	20	95.2	329	2	T30513	hypothetical prote	491	20	95.2	361	2	AG0346	putative ABC-trans
419	20	95.2	329	2	C96033	probable regulator	492	20	95.2	362	2	T19285	hypothetical prote
420	20	95.2	331	2	D69490	LSU ribosomal prot	493	20	95.2	363	2	T40797	60S ribosomal prot
421	20	95.2	331	2	D49905	protein secretion	494	20	95.2	363	2	H87277	flagellin modifika
422	20	95.2	331	2	E83457	molybdopterin bios	495	20	95.2	364	1	QQBE9	BOHRF1 protein - hu
423	20	95.2	331	2	A81173	conserved hypothet	496	20	95.2	364	2	T34954	probable UDP-N-ace
424	20	95.2	331	2	G81932	probable periplasm	497	20	95.2	364	2	T48752	GRRI related prote
425	20	95.2	331	2	T36969	probable phytoene	498	20	95.2	365	1	MXKRAH	nonstructural prot
426	20	95.2	331	2	H90252	NADH dehydrogenase	499	20	95.2	366	2	T50468	probable maturase
427	20	95.2	331	2	A87652	hypothetical prote	500	20	95.2	367	2	AG0543	glutamate 5-kinase
428	20	95.2	332	2	AB0808	probable ion-dhann	501	20	95.2	367	2	KIECEG	glutamate 5-kinase
429	20	95.2	334	1	B49888	thioredoxin-disulf	502	20	95.2	367	2	AG0543	glutamate 5-kinase
430	20	95.2	334	2	A47180	L-lactate dehydrog	503	20	95.2	367	2	C85513	gamma-glutamate k
431	20	95.2	335	2	C69279	heat shock protein	504	20	95.2	367	2	E90662	glutamate 5-kinase
432	20	95.2	335	2	G72304	histidinol-phospha	505	20	95.2	367	2	AE0391	glutamate 5-kinase
433	20	95.2	336	2	G90336	potassium channel	506	20	95.2	368	2	JN0718	cysteine proteinas
434	20	95.2	336	2	B90574	ABC transporter pe	507	20	95.2	369	2	F83250	histidinol-phospha
435	20	95.2	336	2	D98203	probable transcrip	508	20	95.2	370	2	I51199	activin beta B sub
436	20	95.2	337	2	C81801	ketol-acid reducto	509	20	95.2	370	2	A71143	hypothetical prote
437	20	95.2	337	2	F81066	ketol-acid reducto	510	20	95.2	371	2	C83736	transposase (15) B
438	20	95.2	337	2	AF0086	periplasmic bindin	511	20	95.2	371	2	AC3210	hypothetical prote
439	20	95.2	337	2	T33209	hypothetical prote	512	20	95.2	373	2	SI7955	long-chain-fatty-a
440	20	95.2	338	2	B90938	asparaginase (EC 3	513	20	95.2	373	2	S15161	long-chain-fatty-a
441	20	95.2	338	2	AC0711	asparaginase (EC 3	514	20	95.2	374	2	AE2337	hypothetical prote
442	20	95.2	338	2	XDEC1	asparaginase (EC 3	515	20	95.2	374	2	E64489	hypothetical prote
443	20	95.2	338	2	D72327	heat shock opteron	516	20	95.2	375	2	E70011	potassium channel
444	20	95.2	338	2	F87267	PdZ domain family	517	20	95.2	376	2	T35868	probable dipeptida
445	20	95.2	338	2	A84593	hypothetical prote	518	20	95.2	376	2	B75380	probable dipeptida
446	20	95.2	338	2	B96625	protein C14H10.1 (	519	20	95.2	379	2	T10949	cysteine proteinas
447	20	95.2	339	2	AB3330	ketol-acid reducto	520	20	95.2	382	2	T16431	hypothetical prote
448	20	95.2	339	2	D83308	probable transcrip	521	20	95.2	384	2	T21502	hypothetical prote
449	20	95.2	339	2	H83365	hypothetical prote	522	20	95.2	385	2	F97253	UDP-N-acetylglucos
450	20	95.2	340	2	T46112	hypothetical prote	523	20	95.2	385	2	S43540	probable oxidoredu
451	20	95.2	341	1	KIBB36	thymidine kinase (	524	20	95.2	385	2	T36899	ORF MSV214 SCG gen
452	20	95.2	341	1	KIBB40	thymidine kinase (	525	20	95.2	386	2	T28375	hypothetical prote
453	20	95.2	341	1	KIBB73	thymidine kinase (	526	20	95.2	386	2	T38718	hypothetical prote
454	20	95.2	341	1	KIBB73	thymidine kinase (	527	20	95.2	387	2	B83548	hypothetical prote
455	20	95.2	341	1	KIBB73	thymidine kinase (	528	20	95.2	388	2	T26372	DNA topoisomerase
456	20	95.2	341	1	KIBB73	thymidine kinase (	529	20	95.2	388	2	D90248	hypothetical prote
457	20	95.2	342	2	T22428	hypothetical prote	530	20	95.2	389	2	T20604	hypothetical prote
458	20	95.2	343	2	S09777	hypothetical prote	531	20	95.2	389	2	T20604	hypothetical prote
459	20	95.2	343	2	AF3292	chemotaxis motB pr	532	20	95.2	391	2	T25211	probable transcrip
460	20	95.2	343	2	B98267	hypothetical prote	533	20	95.2	391	2	T25211	hypothetical prote
461	20	95.2	343	2	T20388	hypothetical prote	534	20	95.2	392	2	A05025	NADH2 dehydrogenas
462	20	95.2	344	2	S73555	MG415 homolog C12	535	20	95.2	393	2	E95261	serine proteinase
463	20	95.2	345	2	B42604	ORF2 complementary	536	20	95.2	394	2	E64045	mannonate dehydrat
464	20	95.2	345	2	T09018	probable calcium-b	537	20	95.2	394	2	A12546	hypothetical prote
465	20	95.2	345	2	B87375	hypothetical prote	538	20	95.2	394	2	T37520	hypothetical prote
466	20	95.2	345	2	T41473	probable zinc fing	539	20	95.2	394	2	T20519	hypothetical prote
467	20	95.2	345	2	A12560	hypothetical prote	540	20	95.2	395	1	Q0ECTR	hypothetical 45.2K

541	20	95.2	395	2	A91129	hypothetical prote	614	20	95.2	430	2	G72201	conserved hypothet
542	20	95.2	395	2	H85973	hypothetical prote	615	20	95.2	431	2	AE2090	site-specific DNA-
543	20	95.2	395	2	E90896	probable transport	616	20	95.2	431	2	AE0795	isochorismate synt
544	20	95.2	395	2	B85721	membrane protein Y	617	20	95.2	431	2	H81981	probable tyrosine-
545	20	95.2	395	2	A64908	peptidases homolog	618	20	95.2	431	2	AO8060	probable RNA methy
546	20	95.2	396	2	AH1426	peptidases homolog	619	20	95.2	432	2	T33118	hypothetical prote
547	20	95.2	396	2	AG1800	peptidases homolog	620	20	95.2	432	2	H82139	trigger factor [si
548	20	95.2	397	2	B98127	serine proteinase	621	20	95.2	433	2	B84923	hypothetical prote
549	20	95.2	397	2	S26731	neuro-D4 protein -	622	20	95.2	433	2	B84923	hypothetical prote
550	20	95.2	398	2	D83601	conserved hypothet	623	20	95.2	434	2	A81793	tyrosyl-tRNA synth
551	20	95.2	398	2	C96680	hypothetical prote	624	20	95.2	435	2	F98302	edta monooxygenase
552	20	95.2	399	2	S07630	hypothetical prote	625	20	95.2	435	2	A12980	nitrilotriacetate
553	20	95.2	399	2	T10314	viral transcriptio	626	20	95.2	435	2	B83958	glucose-inhibited
554	20	95.2	399	2	T41786	P47 orf40 - Bombyx	627	20	95.2	435	2	F89898	glucose-inhibited
555	20	95.2	400	2	AB1427	carboxypeptidase h	628	20	95.2	436	2	C89926	hypothetical prote
556	20	95.2	400	2	A11800	carboxypeptidase h	629	20	95.2	437	2	T44520	lipopolysaccharide
557	20	95.2	400	2	G69370	3-ketoacyl-CoA thi	630	20	95.2	437	2	T44509	Vi polysaccharide
558	20	95.2	401	2	H72854	viral transcriptio	631	20	95.2	438	2	A86502	ATP synthase subun
559	20	95.2	401	2	F95922	probable NDP-hexos	632	20	95.2	438	2	C72121	ATP synthase, chai
560	20	95.2	403	2	A28451	beta-galactoside a	633	20	95.2	438	2	A83823	ATP-dependent RNA
561	20	95.2	403	2	T39697	DNAJ protein - fis	634	20	95.2	438	2	B97712	hypothetical prote
562	20	95.2	404	2	A99277	hypothetical prote	635	20	95.2	439	2	B98024	homoserine dehydro
563	20	95.2	405	2	AF2559	hypothetical prote	636	20	95.2	439	2	AF0742	conserved hypothet
564	20	95.2	405	2	AE2534	hypothetical prote	637	20	95.2	440	2	D83089	hisidinol dehydro
565	20	95.2	406	2	T02291	hypothetical prote	638	20	95.2	440	2	A72405	DNA repair protein
566	20	95.2	407	1	A40150	inhibin beta-B cha	639	20	95.2	440	2	B64090	dicarboxylate tran
567	20	95.2	407	1	R5PFL1	ribosomal protein	640	20	95.2	440	2	D71715	hypothetical prote
568	20	95.2	407	2	T12085	reverse transcript	641	20	95.2	441	2	G97126	probable Fe-S oxid
569	20	95.2	407	2	T45965	hypothetical prote	642	20	95.2	441	2	G64492	hypothetical prote
570	20	95.2	409	2	F90825	probable integrase	643	20	95.2	442	2	T33412	hypothetical prote
571	20	95.2	410	2	S15163	probable transposa	644	20	95.2	445	2	H96771	hypothetical prote
572	20	95.2	410	2	C44490	retrovirus-related	645	20	95.2	446	2	S46786	serine-tRNA ligase
573	20	95.2	411	2	T47926	hypothetical prote	646	20	95.2	446	2	B70543	hypothetical prote
574	20	95.2	411	2	B41398	inhibin beta-B cha	647	20	95.2	448	2	G89933	hypothetical prote
575	20	95.2	411	2	E64088	ATP-dependent clp	648	20	95.2	448	2	E96837	unknown protein T2
576	20	95.2	412	2	T47321	hypothetical prote	649	20	95.2	449	2	A71248	amidophosphoribosy
577	20	95.2	414	2	AG2407	site-specific DNA-	650	20	95.2	451	2	S76813	hypothetical prote
578	20	95.2	414	2	S60190	vicillin - Zambia fu	651	20	95.2	453	2	T32707	probable initiator
579	20	95.2	416	2	F70017	probable tranfamin	652	20	95.2	454	2	T29024	hypothetical prote
580	20	95.2	416	2	A85684	probable integrase	653	20	95.2	454	2	T44878	3-oxoacyl-lacyl-ca
581	20	95.2	417	2	AF0251	probable M23/M37 p	654	20	95.2	454	2	T41020	probable Uridine k
582	20	95.2	417	2	T47616	hypothetical prote	655	20	95.2	454	2	F70938	probable fabG4 pro
583	20	95.2	417	2	B83469	probable chemotaxi	656	20	95.2	455	2	S16559	cellulase [EC 3.2.
584	20	95.2	418	2	A11784	UV-damage repair p	657	20	95.2	455	2	D75043	seryl-tRNA synthet
585	20	95.2	418	2	AC1409	UV-damage repair p	658	20	95.2	455	2	E71569	probable acyltrans
586	20	95.2	418	2	G83654	transposase (II) B	659	20	95.2	455	2	C97256	uncharacterized pr
587	20	95.2	418	2	H83677	transposase (II) B	660	20	95.2	456	1	E42594	hypothetical prote
588	20	95.2	419	2	T47443	hypothetical prote	661	20	95.2	456	2	T06589	3-methyl-2-oxobuta
589	20	95.2	419	2	A70814	probable integral	662	20	95.2	456	2	T06136	aspartate transami
590	20	95.2	421	2	T51809	succinate-CoA liga	663	20	95.2	460	2	G71117	serine-tRNA ligase
591	20	95.2	423	2	F75635	hypothetical prote	664	20	95.2	460	2	C83468	conserved hypothet
592	20	95.2	423	2	S76384	hypothetical prote	665	20	95.2	461	2	I49366	Gl/S transition co
593	20	95.2	423	2	G83573	conserved hypothet	666	20	95.2	461	2	A57463	p55PIK - mouse
594	20	95.2	423	2	G82145	conserved hypothet	667	20	95.2	462	1	S76100	hypothetical prote
595	20	95.2	423	2	E81010	hemK protein NMB20	668	20	95.2	462	2	T30164	probable phosphor
596	20	95.2	424	2	JC7558	chromatin assembly	669	20	95.2	462	2	AF0057	probable transport
597	20	95.2	424	2	AO2622	conserved hypothet	670	20	95.2	462	2	G01804	interleukin 3-regu
598	20	95.2	425	2	AH1041	Vi polysaccharide	671	20	95.2	463	2	T43344	nuclear receptor N
599	20	95.2	425	2	B36892	Vi polysaccharide	672	20	95.2	465	2	B81658	2-oxo acid dehydro
600	20	95.2	425	2	S36112	Gl/S transition co	673	20	95.2	465	2	E71500	probable lipamide
601	20	95.2	425	2	I39181	Gl/S transition co	674	20	95.2	466	2	AC2696	two component sens
602	20	95.2	426	2	I49367	Gl/S transition co	675	20	95.2	466	2	C97478	popQ protein [impo
603	20	95.2	426	2	B87078	tyrosyl-tRNA synth	676	20	95.2	467	2	S58233	PopQ protein - Rhi
604	20	95.2	426	2	F71299	serine-tRNA ligase	677	20	95.2	468	2	T20475	hypothetical prote
605	20	95.2	427	2	A85789	hypothetical prote	678	20	95.2	471	2	F72337	hypothetical prote
606	20	95.2	427	2	E90940	hypothetical prote	679	20	95.2	471	2	E86322	hypothetical prote
607	20	95.2	427	2	H64938	hypothetical prote	680	20	95.2	471	2	S61202	probable membrane
608	20	95.2	427	2	G84375	adenosylhomocyste	681	20	95.2	472	2	C70853	hypothetical prote
609	20	95.2	428	2	B95158	homoserine dehydro	682	20	95.2	472	2	AH2925	hypothetical prote
610	20	95.2	428	2	B64081	fucose permease ho	683	20	95.2	472	2	F98356	hypothetical prote
611	20	95.2	428	2	AE0712	conserved hypothet	684	20	95.2	473	2	T04225	hypothetical prote
612	20	95.2	429	2	B71307	conserved hypothet	685	20	95.2	473	2	C96516	F16N3.15 [imported
613	20	95.2	430	2	G86870	histidine-tRNA lig	686	20	95.2	474	2	AB2990	polysaccharide bio

687	20	95.2	474	2	G98293	polysaccharid bio	760	20	95.2	535	2	H83324	probable chemotaxi
688	20	95.2	476	2	H84291	hypothetical prote	761	20	95.2	535	2	B45270	sensory histidine
689	20	95.2	478	2	F82175	conserved hypothet	762	20	95.2	535	2	A83202	hypothetical prote
690	20	95.2	479	2	F75113	probable ferredoxi	763	20	95.2	535	2	A87570	ABC transporter, A
691	20	95.2	479	2	A25052	fibrinogen beta ch	764	20	95.2	536	2	T03034	cytochrome p450 -
692	20	95.2	480	1	IBEG	hydroxymethylbilan	765	20	95.2	536	2	T37544	hypothetical serin
693	20	95.2	480	2	A37244	nuclear autoantige	766	20	95.2	537	2	B38179	rpN-2 protein - B
694	20	95.2	480	2	B70446	hypothetical prote	767	20	95.2	537	2	D89889	conserved hypothet
695	20	95.2	481	1	B43674	protein kinase (EC	768	20	95.2	539	2	T36660	hypothetical prote
696	20	95.2	481	1	TVBE17	protein kinase (EC	769	20	95.2	540	2	T47858	proteinnase - Strep
697	20	95.2	481	2	JE0377	p70 S6 kinase (EC	770	20	95.2	541	2	D96779	hypothetical prote
698	20	95.2	482	2	T22981	hypothetical prote	771	20	95.2	541	2	W2BEM5	probable 3-ketoacy
699	20	95.2	485	2	S52411	ZNf165 protein - h	772	20	95.2	543	1	T42932	gene 19 protein -
700	20	95.2	486	2	AI1942	cysteinyI-CRNA syn	773	20	95.2	544	2	T42932	virion tegument pr
701	20	95.2	486	2	AD0373	probable exported	774	20	95.2	544	2	A81178	transport protein
702	20	95.2	487	2	B95275	probable ABC trans	775	20	95.2	545	1	JQ0153	mercury(II) reduct
703	20	95.2	488	2	E95382	hypothetical prote	776	20	95.2	545	2	H83079	hypothetical prote
704	20	95.2	488	2	T10283	ecdysteroid UDPglu	777	20	95.2	545	2	B82740	two-component syst
705	20	95.2	489	2	S23410	FUN19 protein - ye	778	20	95.2	546	2	G72210	hypothetical prote
706	20	95.2	489	2	B72518	hypothetical prote	779	20	95.2	547	2	B86723	NADH oxidase noxC
707	20	95.2	492	2	B70617	hypothetical prote	780	20	95.2	548	2	AB0365	probable pyridine
708	20	95.2	492	2	T10054	pyruvate kinase (E	781	20	95.2	551	2	T40767	hypothetical prote
709	20	95.2	493	2	T74625	NADH-glutamate syn	782	20	95.2	552	2	S55026	secretion protein
710	20	95.2	495	2	S60175	regulatory protein	783	20	95.2	553	1	B55483	transcription init
711	20	95.2	495	2	A48370	nitrogen fixation	784	20	95.2	553	2	S28713	hypothetical prote
712	20	95.2	496	2	C83617	probable aldehyde	785	20	95.2	553	2	AF2662	two component sens
713	20	95.2	497	2	G81598	serine hydroxymeth	786	20	95.2	554	2	A11829	hypothetical prote
714	20	95.2	497	2	C81672	glycine hydroxymet	787	20	95.2	554	2	E90601	hypothetical prote
715	20	95.2	497	2	H71516	probable ABC trans	788	20	95.2	555	2	D90354	NADH oxidase SSO19
716	20	95.2	497	2	G95398	virion proteid hom	789	20	95.2	555	2	B87670	GGDEF family prote
717	20	95.2	504	2	JC1306	probable serine-th	790	20	95.2	555	2	E97444	probable two-compo
718	20	95.2	504	2	T38226	exodeoxyribonuclea	791	20	95.2	555	2	T45351	ferredoxin [import
719	20	95.2	505	2	E87527	hypothetical prote	792	20	95.2	557	2	T24538	hypothetical prote
720	20	95.2	506	2	F83545	glycerol-3-phospha	793	20	95.2	558	2	T32028	hypothetical prote
721	20	95.2	507	2	H82580	histidyl-trna synt	794	20	95.2	558	2	E83905	hypothetical prote
722	20	95.2	507	2	D97441	histidyl-trna synt	795	20	95.2	562	2	S46281	p element - fruit
723	20	95.2	507	2	AF2659	hypothetical prote	796	20	95.2	564	1	HMIVF7	hemagglutinin prec
724	20	95.2	507	2	T19067	hypothetical prote	797	20	95.2	564	2	S15962	hypothetical prote
725	20	95.2	508	1	S46008	probable carbotype	798	20	95.2	564	2	T15477	hypothetical prote
726	20	95.2	509	2	T01344	hypothetical prote	799	20	95.2	565	2	T42593	exonuclease (EC 3.
727	20	95.2	510	2	T07119	cytochrome p450 CP	800	20	95.2	567	2	T30799	hypothetical prote
728	20	95.2	511	1	E89775	2',3'-cyclic-nucle	801	20	95.2	567	2	E72156	C6R protein - vari
729	20	95.2	513	2	C86897	hypothetical prote	802	20	95.2	567	2	E35928	hypothetical prote
730	20	95.2	513	2	S54590	2'-O-ribosyl phosp	803	20	95.2	567	2	I36841	B6R protein - vari
731	20	95.2	513	2	A87324	hypothetical prote	804	20	95.2	567	2	T28485	hypothetical prote
732	20	95.2	514	2	S50785	RRN7 protein - yea	805	20	95.2	568	2	AB3341	precorrin-3B C17-m
733	20	95.2	514	2	D84584	probable protein p	806	20	95.2	569	2	T01399	hypothetical prote
734	20	95.2	514	2	T21286	hypothetical prote	807	20	95.2	569	2	A43317	germ cell-less pro
735	20	95.2	515	2	AF3524	hypothetical prote	808	20	95.2	570	2	C89885	DNA-dependent DNA
736	20	95.2	518	1	A27705	alpha-amylase (EC	809	20	95.2	571	2	H97333	site-specific modi
737	20	95.2	518	2	T43173	myosin heavy chain	810	20	95.2	572	2	B84958	proline-tRNA ligas
738	20	95.2	518	2	D83787	spore germination	811	20	95.2	573	2	A82231	transport ATP-bind
739	20	95.2	518	2	AG0784	rtn protein [import	812	20	95.2	574	2	B29677	complement C9 prec
740	20	95.2	519	2	E86555	serine hydroxymeth	813	20	95.2	574	2	T01131	hypothetical prote
741	20	95.2	519	2	H72067	glycine hydroxymet	814	20	95.2	574	2	B87619	sensor histidine k
742	20	95.2	519	2	T46241	hypothetical prote	815	20	95.2	574	2	S22596	hypothetical prote
743	20	95.2	519	2	T33616	hypothetical prote	816	20	95.2	575	2	T34280	hypothetical prote
744	20	95.2	520	2	H87621	hypothetical prote	817	20	95.2	576	2	A40688	peroxisomal protei
745	20	95.2	521	2	C84249	MAPK-activated pro	818	20	95.2	578	2	H82872	hypothetical prote
746	20	95.2	521	2	T37504	hypothetical prote	819	20	95.2	579	2	F70000	two-component sens
747	20	95.2	522	2	A33644	signal recognition	820	20	95.2	579	2	T16237	hypothetical prote
748	20	95.2	524	2	F75264	hypothetical prote	821	20	95.2	582	2	A43412	semenogelin II pre
749	20	95.2	529	1	W7AD22	early E2A DNA-bind	822	20	95.2	583	1	T10051	pyruvate kinase (E
750	20	95.2	529	1	W7AD22	early E2A DNA-bind	823	20	95.2	584	1	S65587	ABC transporter st
751	20	95.2	529	2	I49504	Tum-P91A antigen -	824	20	95.2	584	2	S51882	topoisomerase I-re
752	20	95.2	529	2	F82983	conserved hypothet	825	20	95.2	585	2	C82157	hypothetical prote
753	20	95.2	530	2	AC2085	phosphodiesterase/	826	20	95.2	585	2	B70503	probable pyrG prot
754	20	95.2	530	2	T48117	hypothetical prote	827	20	95.2	586	2	E84808	hypothetical prote
755	20	95.2	530	2	D70476	DNA helicase - Aqu	828	20	95.2	587	2	T41653	probable transcrip
756	20	95.2	532	2	B82354	deoxycytidylate de	829	20	95.2	589	2	C87664	methyl-accepting c
757	20	95.2	533	2	S33701	homeotic protein D	830	20	95.2	590	2	S72961	CIP synthase (EC 6
758	20	95.2	534	2	JC5096	transposase - fung	831	20	95.2	590	2	E72015	DNA primase - Chla
759	20	95.2	535	2	E96730	hypothetical prote	832	20	95.2	590	2	D86608	DNA primase [import

833	20	95.2	591	2	H86267	probable protein p	906	20	95.2	653	2	B75105	probable DNA helic
834	20	95.2	593	2	S51946	pyruvate kinase (E	907	20	95.2	656	1	QBET2	UL25 protein - hum
835	20	95.2	593	2	A64075	DNA primase (EC 2.	908	20	95.2	656	2	E71080	probable DNA-bind
836	20	95.2	594	2	A81525	DNA primase, proba	909	20	95.2	658	2	E83480	bo-type ubiquinol
837	20	95.2	599	2	C86161	hypothetical prote	910	20	95.2	658	2	D81099	membrane-bound lyc
838	20	95.2	600	2	C95113	oligoendopeptidase	911	20	95.2	658	2	D81842	probable membrane
839	20	95.2	600	2	B97982	group B oligopepti	912	20	95.2	659	2	AF3489	cytochrome o ubiq
840	20	95.2	601	2	AH1003	glutathione-regula	913	20	95.2	660	2	T41580	probable dna-bind
841	20	95.2	601	2	F87548	hypothetical prote	914	20	95.2	660	2	T21551	hypothetical prote
842	20	95.2	605	2	T43974	hypothetical prote	915	20	95.2	661	2	S44773	C29B4.5 protein -
843	20	95.2	608	2	A97685	hypothetical prote	916	20	95.2	663	2	H64312	probable DNA helic
844	20	95.2	608	2	AC2910	hypothetical prote	917	20	95.2	664	2	B36885	ba-type ubiquinol
845	20	95.2	610	2	S19461	probable membrane	918	20	95.2	665	2	A11947	exonuclease ABC c
846	20	95.2	610	2	T44161	hypothetical prote	919	20	95.2	666	2	A42296	lysosome 2 (EC 3.2
847	20	95.2	612	2	S55084	probable membrane	920	20	95.2	666	2	T05432	hypothetical prote
848	20	95.2	614	2	S77221	protein kinase pkn	921	20	95.2	669	2	S74391	exonuclease ABC c
849	20	95.2	616	2	T38717	probable GTP-bind	922	20	95.2	669	2	A47302	proline dehydrogen
850	20	95.2	617	2	G64972	yegA protein precu	923	20	95.2	672	2	G82687	bo-type ubiquinol
851	20	95.2	617	2	B85833	suppressor of ompF	924	20	95.2	672	2	G69503	signal-transducing
852	20	95.2	617	2	G90987	suppressor of ompF	925	20	95.2	676	2	T01084	hypothetical prote
853	20	95.2	618	2	T04237	hypothetical prote	926	20	95.2	676	2	C72749	probable cleavage
854	20	95.2	618	2	AE0770	probable outer mem	927	20	95.2	677	2	T27127	hypothetical prote
855	20	95.2	619	2	C83168	heat shock protein	928	20	95.2	680	2	T04647	hypothetical prote
856	20	95.2	619	2	T44285	heat-shock-cognate	929	20	95.2	681	2	S37809	DNA polymerase sub
857	20	95.2	619	2	B71559	probable metallopr	930	20	95.2	681	2	S83455	transketolase (EC
858	20	95.2	619	2	H81712	conserved hypotet	931	20	95.2	681	2	AF0697	probable type III
859	20	95.2	621	2	S49020	nuclear lamin C pr	932	20	95.2	681	2	T15590	hypothetical prote
860	20	95.2	621	2	F86533	metalloproteinase	933	20	95.2	683	2	S34700	probable purine nu
861	20	95.2	621	2	A72091	metalloproteinase	934	20	95.2	684	2	T47694	probable serine/th
862	20	95.2	621	2	D81578	zinc proteinase CP	935	20	95.2	686	2	D96611	probable CRK1 prot
863	20	95.2	621	2	AB3317	ABC transporter AT	936	20	95.2	687	2	E81027	glycyl-tRNA synth
864	20	95.2	623	2	T35377	probable membrane	937	20	95.2	687	2	G81970	probable glycine-t
865	20	95.2	624	1	VGNV87	p87 capsid protein	938	20	95.2	688	2	E86409	hypothetical prote
866	20	95.2	624	2	T10374	p87 capsid protein	939	20	95.2	688	2	T32750	hypothetical prote
867	20	95.2	625	2	C25977	phosphotransferase	940	20	95.2	688	2	AB2379	two-component sens
868	20	95.2	625	2	S60401	nuclear division p	941	20	95.2	689	1	JW0107	very-long-chain ac
869	20	95.2	626	1	S77286	exonuclease ABC c	942	20	95.2	691	2	T46476	hypothetical prote
870	20	95.2	628	2	A81999	glucose inhibited	943	20	95.2	693	2	T26415	hypothetical prote
871	20	95.2	631	2	F81227	glucose inhibited	944	20	95.2	697	2	D84429	hypothetical prote
872	20	95.2	631	2	D95348	nitric-oxide reduc	945	20	95.2	697	2	F84646	hypothetical prote
873	20	95.2	632	2	T38617	probable ubiquitin	946	20	95.2	699	2	T16109	hypothetical prote
874	20	95.2	632	2	AF3095	nitric oxide reduc	947	20	95.2	703	2	T05632	hypothetical prote
875	20	95.2	632	2	D98191	trkA-like protein	948	20	95.2	704	2	B84530	probable RING zinc
876	20	95.2	633	2	T14612	hypothetical prote	949	20	95.2	705	2	D75345	hypothetical prote
877	20	95.2	633	2	T28788	hypothetical prote	950	20	95.2	705	2	T51788	hypothetical prote
878	20	95.2	634	2	A83447	heat shock protein	951	20	95.2	705	2	A48144	protein kinase CDC
879	20	95.2	634	2	T32324	hypothetical prote	952	20	95.2	706	2	T36176	hypothetical prote
880	20	95.2	637	2	A82738	heat shock protein	953	20	95.2	706	2	T41024	hypothetical prote
881	20	95.2	637	2	D87335	DNA mismatch repai	954	20	95.2	708	2	T22377	hypothetical prote
882	20	95.2	638	2	B83890	hypothetical prote	955	20	95.2	709	2	A98299	3-methylcrotonyl-C
883	20	95.2	639	2	A60633	tetracycline resis	956	20	95.2	714	1	S44198	DNA topoisomerase
884	20	95.2	640	2	T26820	hypothetical prote	957	20	95.2	715	2	A34408	peroxidase (EC 1.1
885	20	95.2	642	2	A11827	cyclomaltodextrin	958	20	95.2	717	2	T29816	hypothetical prote
886	20	95.2	642	2	D69085	transcription cont	959	20	95.2	719	2	T02154	protein kinase hom
887	20	95.2	642	2	AF2995	conserved hypotet	960	20	95.2	719	2	S51739	transcription repr
888	20	95.2	643	2	A43423	dynein 74k chain,	961	20	95.2	719	2	A81358	hypothetical prote
889	20	95.2	644	2	B85758	RNase II, mRNA deg	962	20	95.2	720	2	C84055	cell division prot
890	20	95.2	644	2	AB0656	exoribonuclease II	963	20	95.2	724	2	D85075	probable athila tr
891	20	95.2	644	2	A64877	exoribonuclease II	964	20	95.2	725	2	T21363	hypothetical prote
892	20	95.2	644	2	C90861	RNase II, mRNA deg	965	20	95.2	725	2	S64124	probable membrane
893	20	95.2	644	2	AF0272	exoribonuclease II	966	20	95.2	729	2	AF3299	malate synthase (E
894	20	95.2	645	2	H86281	protein F10B6.18 l	967	20	95.2	729	2	T23474	hypothetical prote
895	20	95.2	645	2	D90782	probable tail fi	968	20	95.2	730	2	D87365	prolyl oligopeptid
896	20	95.2	645	2	H85642	probable tail fi	969	20	95.2	733	2	E95335	probable cation tr
897	20	95.2	646	2	T34532	hypothetical prote	970	20	95.2	733	2	D83588	conserved hypotet
898	20	95.2	647	2	T49586	related to nif-spe	971	20	95.2	733	2	F84476	probable Athila re
899	20	95.2	648	2	C69423	DNA helicase homol	972	20	95.2	735	2	T40619	neural trehalase
900	20	95.2	650	2	G72429	hypothetical prote	973	20	95.2	738	2	T38767	probable RNA-bind
901	20	95.2	650	2	D1021	hypothetical prote	974	20	95.2	741	2	AF2297	hypothetical prote
902	20	95.2	650	2	D98288	tfua (U39409) (imp	975	20	95.2	743	2	T26102	hypothetical prote
903	20	95.2	650	2	F70974	probable acral pro	976	20	95.2	743	2	E87386	periplasmic beta-g
904	20	95.2	652	2	B82724	cardiolipin syntha	977	20	95.2	744	2	AH2582	malate synthase G
905	20	95.2	652	2	AD2316	hypothetical prote	978	20	95.2	744	2	F97364	malate synthase G

979 20 95.2 747 2 F64127 DNA topoisomerase  
980 20 95.2 747 2 S01918 sex-specific stora  
981 20 95.2 749 2 E95889 probable dehydroge  
982 20 95.2 749 2 A45046 basic juvenile hor  
983 20 95.2 749 2 C84508 probable cap-bind  
984 20 95.2 751 2 C84367 DNA binding protei  
985 20 95.2 751 2 T15230 hypothetical prote  
986 20 95.2 752 2 AF0888 topoisomerase IV c  
987 20 95.2 752 2 A65089 DNA topoisomerase  
988 20 95.2 752 2 A45582 DNA topoisomerase  
989 20 95.2 752 2 G85961 DNA topoisomerase  
990 20 95.2 752 2 G91116 DNA topoisomerase  
991 20 95.2 753 2 S22802 transposase (clone  
992 20 95.2 753 2 T01619 hypothetical proce  
993 20 95.2 757 2 A48592 transferrin recept  
994 20 95.2 760 2 T31556 hypothetical prote  
995 20 95.2 760 2 G86373 protein T23E23.14  
996 20 95.2 761 2 F82078 topoisomerase IV,  
997 20 95.2 761 2 T43788 DNA topoisomerase  
998 20 95.2 765 2 T49592 neurofilament trip  
999 20 95.2 766 2 A24786 P element transpos  
1000 20 95.2 766 2 H83141 probable two-compo

## ALIGNMENTS

RESULT 1  
AF0749  
Hypothetical protein STY2156 [imported] - Salmonella enterica subsp. enterica serovar Typhimurium  
C:Species: Salmonella enterica subsp. enterica serovar typhi  
A:Note: this species has also been called Salmonella typhi  
C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
C:Accession: AF0749  
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, F.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.  
Nature 413, 848-852, 2001  
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; White, N.; Farrar, S.; O'Gaora, P.  
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhimurium  
A:Reference number: AB0502; MUID:21534947; PMID:11677608  
A:Accession: AF0749  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-51 <PAR>  
A:Cross-references: UNIPARC:UPI0000059CC5; GB:AL513382; PIDN:CAD05697.1; PID:G16503191; C:Genetics:  
A:Gene: STY2156

Query Match 95.2%; Score 20; DB 2; Length 51;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
Db 33 RRLN 36

RESULT 2  
S00016  
protamine Z1 - smaller spotted catshark  
C:Species: Scyliorhinus canicula (smaller spotted dogfish)  
C:Date: 07-Mar-1988 #sequence\_revision 07-Mar-1988 #text\_change 09-Jul-2004  
C:Accession: S00016; S01463  
R:Berlot-Picard, P.; Vlodjani, G.; Doly, J.  
Eur. J. Biochem. 165, 553-557, 1987  
A:Title: Isolation and characterization of a cDNA clone encoding testis protamine Z1 from the Atlantic spotted catshark  
A:Reference number: S00016; MUID:87246639; PMID:3595601  
A:Accession: S00016  
A:Molecule type: mRNA  
A:Residues: 1-51 <BER>  
A:Cross-references: UNIPROT:P08433; UNIPARC:UPI0000171494; GB:X05611; NID:G64302; PIDN:C100000004; R:Gusse, M.; Sautiere, P.; Chauviere, M.; Chevallier, P.

Biochim. Biophys. Acta 748, 93-98, 1983  
A:Title: Extraction, purification and characterization of the sperm protamines of the dogfish  
A:Reference number: S01463; MUID:84000513; PMID:6615852  
A:Accession: S01463  
A:Molecule type: protein  
A:Residues: 2-6 <GUS>  
A:Cross-references: UNIPARC:UPI000017BF3D  
C:Keywords: DNA binding; sperm; testis  
F:2-51/Product: protamine Z1 #status predicted <MAT>

Query Match 95.2%; Score 20; DB 2; Length 51;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
Db 21 RRLN 24

RESULT 3  
D98179  
Hypothetical protein AGR\_L\_771 [imported] - Agrobacterium tumefaciens (strain C58, Cerevisiae)  
C:Species: Agrobacterium tumefaciens  
C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
C:Accession: D98179  
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.  
Science 294, 2323-2328, 2001  
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens  
A:Reference number: A97359; MUID:21608551; PMID:11743194  
A:Accession: D98179  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-51 <KUR>  
A:Cross-references: UNIPROT:Q8U429; UNIPARC:UPI00000D2911; GB:AE007870; PIDN:AAK88958.1  
C:Genetics:  
A:Gene: AGR\_L\_771  
A:Map position: linear chromosome

Query Match 95.2%; Score 20; DB 2; Length 51;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
Db 34 RRLN 37

RESULT 4  
B84091  
Hypothetical protein B84091 [imported] - Bacillus halodurans (strain C-125)  
C:Species: Bacillus halodurans  
C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 09-Jul-2004  
C:Accession: B84091  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirai, T.  
Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and its relationship to other members of the family Halobacteriaceae  
A:Reference number: A83650; MUID:20512582; PMID:11058132  
A:Accession: B84091  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-61 <STO>  
A:Cross-references: UNIPROT:Q9K743; UNIPARC:UPI00000C421D; GB:AP001519; GB:BA000004; R:Gusse, M.; Sautiere, P.; Chauviere, M.; Chevallier, P.

Query Match 95.2%; Score 20; DB 2; Length 61;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
Db 1 RRLN 4



Db 11 RRLN 14

RESULT 5  
B64479  
hypothetical protein MJ1435 - Methanococcus jannaschii  
C:Species: Methanococcus jannaschii  
C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 09-Jul-2004  
C:Accession: B64479  
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, R.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; Ison, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Huret, M.A.  
Science 273, 1058-1073, 1996  
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.  
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii  
A:Reference number: A64300; MUID:96337999; PMID:8688087  
A:Accession: B64479  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-71 <BUL>  
A:Cross-references: UNIPROT:Q58830; UNIPARC:UPI000013AACB; GB:U67584; GB:L77117; NID:gl5  
C:Genetics:  
A:Map position: FOR1404404-1404619

Query Match 95.2%; Score 20; DB 2; Length 71;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
|||||  
Db 21 RRLN 24

RESULT 6  
I39470  
apolipoprotein B - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 06-Sep-1996 #sequence\_revision 06-Sep-1996 #text\_change 21-Jul-2000  
C:Accession: I39470  
R:Huang, L.S.; Rippes, M.E.; Korman, S.H.; Deckelbaum, R.J.; Breslow, J.L.  
J. Biol. Chem. 264, 11394-11400, 1989  
A:Title: Hypobetalipoproteinemia due to an apolipoprotein B gene exon 21 deletion derivative  
A:Reference number: I39469; MUID:89231895; PMID:2567736  
A:Accession: I39470  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-72 <RES>  
A:Cross-references: UNIPARC:UPI000011P7BE; GB:J04838; NID:gl78737; PIDN:AAAS3374.1; PID:  
C:Genetics:  
A:Gene: GDB:APOB  
A:Cross-references: GDB:119686; OMIM:107730  
A:Map position: 2p24-2p24  
A:Introns: 58/3

Query Match 95.2%; Score 20; DB 2; Length 72;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
|||||  
Db 61 RRLN 64

RESULT 7  
E90729  
hypothetical protein ECs0805 [imported] - Escherichia coli (strain O157:H7, substrain R1)  
C:Species: Escherichia coli  
C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004  
C:Accession: E90729  
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen

A:Reference number: A99629; MUID:21156231; PMID:11258796  
A:Accession: E90729  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-73 <HAY>  
A:Cross-references: UNIPROT:Q8X3F3; UNIPARC:UPI00000D2997; GB:BA000007; PIDN:BAB34228.1  
A:Experimental source: strain O157:H7, substrain R1MD 0509952  
C:Genetics:  
A:Gene: ECs0805

Query Match 95.2%; Score 20; DB 2; Length 73;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
|||||  
Db 23 RRLN 26

RESULT 8  
B90775  
C4-type zinc finger protein (TraR family) [imported] - Escherichia coli (strain O157:H7)  
C:Species: Escherichia coli  
C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004  
C:Accession: B90775  
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen  
A:Reference number: A99629; MUID:21156231; PMID:11258796  
A:Accession: B90775  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-73 <HAY>  
A:Cross-references: UNIPROT:Q9KXG7; UNIPARC:UPI000009C1ED; GB:BA000007; PIDN:BAB34593.1  
A:Experimental source: strain O157:H7, substrain R1MD 0509952  
C:Genetics:  
A:Gene: ECs1170

Query Match 95.2%; Score 20; DB 2; Length 73;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
|||||  
Db 23 RRLN 26

RESULT 9  
F91004  
probable C4-type zinc finger protein [imported] - Escherichia coli (strain O157:H7, sub  
C:Species: Escherichia coli  
C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004  
C:Accession: F91004  
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen  
A:Reference number: A99629; MUID:21156231; PMID:11258796  
A:Accession: F91004  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-73 <HAY>  
A:Cross-references: UNIPROT:Q9EYAL; UNIPARC:UPI00000D2950; GB:BA000007; PIDN:BAB36429.1  
A:Experimental source: strain O157:H7, substrain R1MD 0509952  
C:Genetics:  
A:Gene: ECs3006

Query Match 95.2%; Score 20; DB 2; Length 73;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4

Db	23 RRLN 26		
RESULT 10			
C91287	hypothetical protein XF0242 [imported] - Xylella fastidiosa (strain 9a5c)		
C:Species: Xylella fastidiosa			
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004			
C:Accession: E82829			
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing			
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.			
A:Reference number: A82515; MUID:20365717; PMID:10910347			
A:Note: for a complete list of authors see reference number A59328 below			
A:Accession: E82829			
A:Status: Preliminary			
A:Molecule type: DNA			
A:Residues: 1-74 <SIM>			
A:Cross-references: UNIPROT:Q9PG06; UNIPARC:UPI00000C2341; GB:AE003878; GB:AE003849; NID:9			
A:Experimental source: strain 9a5c			
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A			
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H			
as-Neco, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.			
submitted to GenBank, June 2000			
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm			
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig			
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E			
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, B.C.; Miyaki, C.Y.			
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A			
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak			
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv			
M.; Tshuko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z			
A:Reference number: A59328			
A:Contents: annotation			
C:Genetics:			
A:Gene: XF0242			
Query Match 95.2%; Score 20; DB 2; Length 74;			
Best Local Similarity 100.0%; Pred. No. 1.8e+02;			
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY 1 RRLN 4			
Db 10 RRLN 13			
RESULT 11			
C64871	hypothetical protein b1240 - Escherichia coli (strain K-12)		
C:Species: Escherichia coli			
C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004			
C:Accession: C64871			
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co			
.A.; Rose, D.J.; Mau, B.; Shao, Y.			
Science 277, 1453-1462, 1997			
A:Title: The complete genome sequence of Escherichia coli K-12.			
A:Reference number: A64720; MUID:97426617; PMID:9278503			
A:Accession: C64871			
A:Status: nucleic acid sequence not shown; translation not shown			
A:Molecule type: DNA			
A:Residues: 1-76 <BLAT>			
A:Cross-references: UNIPROT:P76024; UNIPARC:UPI00000D35B4; GB:AE000222; GB:U00096; NID:9			
A:Experimental source: strain K-12, substrain MG1655			
Query Match 95.2%; Score 20; DB 2; Length 76;			
Best Local Similarity 100.0%; Pred. No. 1.8e+02;			
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY 1 RRLN 4			
Db 28 RRLN 31			
RESULT 12			
C91287	hypothetical protein ECS267 [imported] - Escherichia coli (strain O157:H7, substrain R1		
C:Species: Escherichia coli			
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004			
C:Accession: C91287			
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.			
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.			
DNA Res. 8, 11-22, 2001			
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen			
A:Reference number: A99629; MUID:21156231; PMID:11258796			
A:Accession: C91287			
A:Status: Preliminary			
A:Molecule type: DNA			
A:Residues: 1-80 <HAY>			
A:Cross-references: UNIPROT:Q8X2N0; UNIPARC:UPI00000D2A99; GB:BA000007; PIDN:BA838690.1.			
A:Experimental source: strain O157:H7, substrain R1MD 0509952			
C:Genetics:			
A:Gene: ECS267			
Query Match 95.2%; Score 20; DB 2; Length 80;			
Best Local Similarity 100.0%; Pred. No. 1.9e+02;			
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY 1 RRLN 4			
Db 31 RRLN 34			
RESULT 13			
T08595	cysteine proteinase (EC 3.4.22.-) - rape (fragment)		
C:Species: Brassica napus (rape)			
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004			
C:Accession: T08595			
R:Buchanan-Wollaston, V.; Ainsworth, C.			
Plant Mol. Biol. 33, 821-834, 1997			
A:Title: Leaf senescence in Brassica napus: cloning of senescence related genes by sub			
A:Reference number: Z16446; MUID:97260386; PMID:9106506			
A:Accession: T08595			
A:Status: translated from GB/EMBL/DBJ			
A:Molecule type: mRNA			
A:Residues: 1-85 <BUC>			
A:Cross-references: UNIPROT:Q96352; UNIPARC:UPI00000A98F0; EMBL:U68221; NID:q1527224; P			
A:Experimental source: cv. Falcon; senescing leaves			
C:Genetics:			
A:Gene: LSC790			
C:Superfamily: papain			
C:Keywords: cysteine proteinase; hydrolase			
F;58,59,60/Binding site: substrate (Asn, Ser, Trp) #status predicted			
Query Match 95.2%; Score 20; DB 2; Length 85;			
Best Local Similarity 100.0%; Pred. No. 2e+02;			
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY 1 RRLN 4			
Db 27 RRLN 30			
RESULT 14			
S00180	spermatid protein S1 - smaller spotted catshark		
C:Species: Scyliorhinus canicula (smaller spotted catshark, smaller spotted dogfish)			
C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 09-Jul-2004			
C:Accession: S00180			
R:Chauviere, M.; Martinage, A.; Briand, G.; Sautiere, P.; Chevaillier, P.			
Eur. J. Biochem. 169, 105-111, 1987			
A:Title: Nuclear basic protein transition during sperm differentiation: amino acid sequ			
A:Reference number: S00180; MUID:88055030; PMID:3678229			
A:Accession: S00180			
A:Molecule type: protein			

A:Residues: 1-87 <CHA>  
A:Cross-references: UNIPROT:P13275; UNIPARC:UPI000013603D  
C:Superfamily: sperm histone  
C:Keywords: DNA binding; nucleus; phosphoprotein; sperm  
F:3.55/Binding site: phosphate (Ser) (covalent) #status experimental

Query Match 95.2%; Score 20; DB 2; Length 87;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
Db 70 RRLN 73

RESULT 15  
B37475  
probable structural component p14 - borna disease virus  
C:Species: borna disease virus  
C:Date: 16-Feb-1994 #sequence\_revision 18-Nov-1994 #text\_change 16-Aug-2004  
C:Accession: B37475  
R:Pyper, J.M.; Richt, J.A.; Brown, L.; Rott, R.; Narayan, O.; Clements, J.E.  
Virology 195, 229-238, 1993  
A:Title: Genomic organization of the structural proteins of borna disease virus revealed  
A:Reference number: A37475; MUID:93303922; PMID:8317098  
A:Accession: B37475  
A:Status: preliminary  
A:Molecule type: nucleic acid  
A:Residues: 1-87 <PVP>  
A:Cross-references: UNIPROT:Q86622; UNIPARC:UPI00000F867D; GB:S62821; NID:g386390; PIDN:  
A:Note: sequence extracted from NCBI backbone (NCBIN:134146, NCBIP:134148)  
C:Superfamily: Borna disease virus 24K antigen

Query Match 95.2%; Score 20; DB 2; Length 87;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
Db 14 RRLN 17

RESULT 16  
C90392  
conserved hypothetical protein [imported] - Sulfolobus solfataricus  
C:Species: Sulfolobus solfataricus  
C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 09-Jul-2004  
C:Accession: C90392  
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awa'ez, M.J.; Chan-  
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, H.  
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.  
submitted to GenBank, April 2001  
A:Description: Sulfolobus solfataricus complete genome.  
A:Reference number: A99139  
A:Accession: C90392  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-89 <KUR>  
A:Cross-references: UNIPROT:Q97WJ2; UNIPARC:UPI00000646A1; GB:AE006641; NID:g13815524; F:  
C:Genetics:  
A:Gene: SSO10051

Query Match 95.2%; Score 20; DB 2; Length 89;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
Db 37 RRLN 40

RESULT 17  
H83003  
hypothetical protein YPO3540 [imported] - Yersinia pestis (strain CO92)  
C:Species: Yersinia pestis  
C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004

conserved hypothetical protein PA5148 [imported] - Pseudomonas aeruginosa (strain PA01)  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 05-Oct-2004  
C:Accession: H83003  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B.  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path  
A:Reference number: A82950; MUID:20437337; PMID:10984043  
A:Accession: H83003  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-90 <STO>  
A:Cross-references: UNIPARC:UPI00000C5F26; GB:AE004927; GB:AE004091; NID:g9951437; PIDN:  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA5148  
C:Superfamily: fe(II) trafficking protein YggX

Query Match 95.2%; Score 20; DB 2; Length 90;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
Db 53 RRLN 56

RESULT 18  
A95415  
hypothetical protein SMA2275 [imported] - Sinorhizobium meliloti (strain 1021) magaplas  
C:Species: Sinorhizobium meliloti  
C:Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 09-Jul-2004  
C:Accession: A95415  
R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bow  
.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.  
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001  
A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilo  
A:Reference number: A95262; MUID:21396509; PMID:11481432  
A:Accession: A95415  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-91 <KUR>  
A:Cross-references: UNIPROT:Q92XL9; UNIPARC:UPI00000CB35D; GB:AE006469; PIDN:AAK65883.1.  
A:Experimental source: strain 1021, megaplasmid pSymA  
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, F.  
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;  
L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001  
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,  
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.  
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.  
A:Reference number: A96039; MUID:21368234; PMID:11474104  
A:Contents: annotation  
C:Genetics:  
A:Gene: SMA2275  
A:Genome: plasmid

Query Match 95.2%; Score 20; DB 2; Length 91;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
Db 60 RRLN 63

RESULT 19  
AE0430  
hypothetical protein YPO3540 [imported] - Yersinia pestis (strain CO92)  
C:Species: Yersinia pestis  
C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004

C;Accession: AE0430  
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Hill, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, Nature 413, 523-527, 2001  
A;Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.  
A;Reference number: AB0001; MUID:21470413; PMID:11586360  
A;Accession: AE0430  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-92 <KUR>  
A;Cross-references: UNIPROT:Q8ZB80; UNIPARC:UPI00000CDA3A; GB:AL590842; PIDN:CAC92769.1;  
C;Genetics:  
A;Gene: YPO3540

Query Match 95.2%; Score 20; DB 2; Length 92;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
|  
|  
|  
|  
Db 87 RRLN 90

## RESULT 20

RBBP22  
abc1 protein - phage P22  
C;Species: phage P22  
A;Note: host *Salmonella typhimurium*  
C;Date: 31-Mar-1989 #sequence\_revision 31-Mar-1989 #text\_change 09-Jul-2004  
C;Accession: A27841  
R;Murphy, K.C.; Fenton, A.C.; Poteete, A.R. Virolgy 160, 456-464, 1987  
A;Title: Sequence of the bacteriophage P22 anti-recBCD (abc) genes and properties of P22  
A;Reference number: A94366; MUID:88019195; PMID:3660589  
A;Accession: A27841  
A;Molecule type: DNA  
A;Residues: 1-94 <MUR>  
A;Cross-references: UNIPROT:P11190; UNIPARC:UPI000003A83; GB:J02471; GB:M17737; NID:921  
A;Note: the authors translated the codon TTA for residues 4 and 91 as Lys, ACG for residue  
C;Genetics:  
A;Gene: abc1  
C;Superfamily: phage P22 abc1 protein

Query Match 95.2%; Score 20; DB 1; Length 94;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
|  
|  
|  
|  
Db 2 RRLN 5

## RESULT 21

AD0229  
hypothetical protein YPO1879 [imported] - *Yersinia pestis* (strain CO92)  
C;Species: *Yersinia pestis*  
C;Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004  
C;Accession: AD0229  
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Hill, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, Nature 413, 523-527, 2001  
A;Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.  
A;Reference number: AB0001; MUID:21470413; PMID:11586360  
A;Accession: AD0229  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-95 <KUR>  
A;Cross-references: UNIPROT:Q8ZF39; UNIPARC:UPI00000DC7AF; GB:AL590842; PIDN:CAC90696.1;  
C;Genetics:  
A;Gene: YPO1879

Query Match 95.2%; Score 20; DB 2; Length 95;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
|  
|  
|  
|  
Db 84 RRLN 87

## RESULT 22

B26074  
cysteine proteinase (EC 3.4.22.-) 13 - papaya (fragment)  
C;Species: *Carica papaya* (papaya)  
C;Date: 05-Oct-1988 #sequence\_revision 05-Oct-1988 #text\_change 09-Jul-2004  
C;Accession: B26074  
R;McKee, R.A.; Adams, S.; Matthews, J.A.; Smith, C.J.; Smith, H. Biochem. J. 237, 105-110, 1986  
A;Title: Molecular cloning of two cysteine proteinases from paw-paw (*Carica papaya*).  
A;Reference number: A90332; MUID:87099799; PMID:3541893  
A;Accession: B26074  
A;Molecule type: mRNA  
A;Residues: 1-96 <MCK>  
A;Cross-references: UNIPROT:P05993; UNIPARC:UPI0000131300; GB:X03971; GB:M25360; NID:91;  
C;Superfamily: papain  
C;Keywords: cysteine proteinase; hydrolase

Query Match 95.2%; Score 20; DB 2; Length 96;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
|  
|  
|  
|  
Db 27 RRLN 30

## RESULT 23

S51929  
homeotic protein CHB5 - carrot  
C;Species: *Daucus carota* (carrot)  
C;Date: 14-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 05-Oct-2004  
C;Accession: S51929  
R;Kawahara, R.; Komamine, A.; Fukuda, H. Plant Mol. Biol. 27, 155-164, 1995  
A;Title: Isolation and characterization of homeobox-containing genes of carrot.  
A;Reference number: S51925; MUID:95169997; PMID:7865785  
A;Accession: S51929  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-96 <KAW>  
A;Cross-references: UNIPARC:UPI000017A2CD  
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation  
F;2-58/Domain: homeobox homology <HGX>

Query Match 95.2%; Score 20; DB 2; Length 96;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
|  
|  
|  
|  
Db 6 RRLN 9

## RESULT 24

S51928  
homeotic protein CHB4 - carrot  
C;Species: *Daucus carota* (carrot)  
C;Date: 14-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 05-Oct-2004  
C;Accession: S51928  
R;Kawahara, R.; Komamine, A.; Fukuda, H. Plant Mol. Biol. 27, 155-164, 1995  
A;Title: Isolation and characterization of homeobox-containing genes of carrot.  
A;Reference number: S51925; MUID:95169997; PMID:7865785  
A;Accession: S51928

A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-96 <KAW>  
A;Cross-references: UNIPROT:Q43428; UNIPARC:UPI000017A2CC  
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation  
F;2-58/Domain: homeobox homology <HOX>

Query Match 95.2%; Score 20; DB 2; Length 96;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
|||||  
Db 6 RRLN 9

RESULT 25  
S62346  
L71-5 protein - fruit fly (*Drosophila melanogaster*)  
C;Species: *Drosophila melanogaster*  
C;Date: 19-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 09-Jul-2004  
C;Accession: S62346  
R;Wright, L.G.; Chen, T.; Thummel, C.S.; Guild, G.M.  
J. Mol. Biol. 255, 387-400, 1996  
A;Title: Molecular characterization of the 71E late puff in *Drosophila melanogaster* reveals  
A;Reference number: S62333; MUID:96152797; PMID:8568894  
A;Accession: S62346  
A;Status: nucleic acid sequence not shown  
A;Molecule type: mRNA  
A;Residues: 1-98 <WRI>  
A;Cross-references: UNIPROT:Q24074; UNIPARC:UPI0000083ECE; EMBL:U24244; NID:g775234; PID  
C;Genetics:  
A;Gene: L71-5  
A;Cross-references: FlyBase:FBgn0004593  
C;Superfamily: L71-10 protein

Query Match 95.2%; Score 20; DB 2; Length 98;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
|||||  
Db 37 RRLN 40

RESULT 26  
A39437  
exopolysaccharide synthesis protein exoX - *Rhizobium meliloti*  
C;Species: *Rhizobium meliloti*  
C;Date: 21-Feb-1992 #sequence\_revision 21-Feb-1992 #text\_change 09-Jul-2004  
C;Accession: A39437  
R;Reed, J.W.; Capage, M.; Walker, G.C.  
J. Bacteriol. 173, 3776-3788, 1991  
A;Title: *Rhizobium meliloti* exoG and exoJ mutations affect the exoX-exoY system for modu  
A;Reference number: A39437; MUID:91267943; PMID:2050634  
A;Accession: A39437  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-98 <REE>  
A;Cross-references: UNIPROT:Q02730; UNIPARC:UPI000012A387; GB:M61751; NID:g152185; PIDN:

Query Match 95.2%; Score 20; DB 2; Length 98;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
|||||  
Db 88 RRLN 91

RESULT 27  
F95975  
posttranscription regulator, repressor protein [imported] - *Sinorhizobium meliloti* (stra

C;Species: *Sinorhizobium meliloti*  
C;Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 09-Jul-2004  
C;Accession: F95975  
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Herna  
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
A;Title: The complete sequence of the 1.683-kb pSymB megaplasmid from the N2-fixing end  
A;Reference number: A95842; MUID:21396508; PMID:11481431  
A;Accession: F95975  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-98 <KUR>  
A;Cross-references: UNIPROT:Q02730; UNIPARC:UPI000012A387; GB:AL591985; PIDN:CAC49470.1  
A;Experimental source: strain 1021, megaplasmid pSymB  
R;Galibert, F.; Finan, T.M.; Long, S.R.; Punler, A.; Abola, P.; Ampe, F.; Barloy-Hubler  
peia, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.  
L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001  
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure  
heault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K  
A;Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.  
A;Reference number: A96039; MUID:21368234; PMID:11474104  
A;Contents: annotation  
C;Genetics:  
A;Gene: exoX; Smb20947  
A;Genome: plasmid

Query Match 95.2%; Score 20; DB 2; Length 98;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
|||||  
Db 88 RRLN 91

RESULT 28  
G96029  
hypothetical protein [imported] - *Sinorhizobium meliloti* (strain 1021) megaplasmid pSym  
C;Species: *Sinorhizobium meliloti*  
C;Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 09-Jul-2004  
C;Accession: G96029  
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Herna  
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
A;Title: The complete sequence of the 1.683-kb pSymB megaplasmid from the N2-fixing end  
A;Reference number: A95842; MUID:21396508; PMID:11481431  
A;Accession: G96029  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-98 <KUR>  
A;Cross-references: UNIPROT:Q92TK8; UNIPARC:UPI00000CB8B2; GB:AL591985; PIDN:CAC49903.1  
R;Galibert, F.; Finan, T.M.; Long, S.R.; Punler, A.; Abola, P.; Ampe, F.; Barloy-Hubler  
peia, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.  
L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001  
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure  
heault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K  
A;Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.  
A;Reference number: A96039; MUID:21368234; PMID:11474104  
A;Contents: annotation  
C;Genetics:  
A;Gene: Smb20591  
A;Genome: plasmid

Query Match 95.2%; Score 20; DB 2; Length 98;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
|||||  
Db 33 RRLN 36

E82425  
 hypotoxin-like protein VCA0721 [imported] - *Vibrio cholerae* (strain N16961 serogroup O1)  
 C:Species: *Vibrio cholerae*  
 C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
 C:Accession: E82425  
 R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;  
 Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, F.  
 1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
 Nature 406, 477-483, 2000  
 #:Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.

## F84479

En/Spm-like transposon protein [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
 C:Accession: F84479  
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;  
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, L.  
 Nature 402, 761-768, 1999  
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A:Reference number: A84420; MUID:20083487; PMID:10617197  
 A:Accession: F84479  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-108 <STO>  
 A:Cross-references: UNIPROT:Q9ZU61; UNIPARC:UPI00000A1DD3; GB:AE002093; NID:g4262212; PIR  
 C:Genetics:  
 A:Gene: At2g06730  
 A:Map position: 2

Query Match 95.2%; Score 20; DB 2; Length 108;  
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
 ||||  
 Db 97 RRLN 100

## RESULT 35

## CABO

cytochrome-c oxidase (BC 1.9.3.1) chain Va [validated] - bovine  
 C:Species: Bos primigenius taurus (cattle)  
 C>Date: 30-Nov-1979 #sequence\_revision 08-Oct-1981 #text\_change 09-Jul-2004  
 C:Accession: A0493; D23968  
 R:Ranaka, M.; Hanlu, M.; Yasunobu, K.T.; Yu, C.A.; Yu, L.; Wei, Y.H.; King, T.E.  
 J. Biol. Chem. 254, 3879-3885, 1979  
 A:Title: Amino acid sequence of subunit V of bovine heart cytochrome oxidase, the heme a  
 A:Reference number: A00493; MUID:79173095; PMID:220224  
 A:Accession: A00493

A:Molecule type: protein  
 A:Residues: 1-109 <TAN>  
 A:Cross-references: UNIPROT:P00426; UNIPARC:UPI0000112899  
 A:Experimental source: heart  
 R:Yanamura, W.; Zhang, Y.Z.; Takamiya, S.; Capaldi, R.A.  
 Biochemistry 27, 4909-4914, 1988  
 A:Title: Tissue-specific differences between heart and liver cytochrome c oxidase.  
 A:Reference number: A90531; MUID:89000697; PMID:2844245

A:Accession: D23968  
 A:Molecule type: protein  
 A:Residues: 1-34 <YAN>  
 A:Cross-references: UNIPARC:UPI000017217B  
 A:Experimental source: liver  
 R:Tsuikihara, T.; Aoyama, H.; Yamashita, E.; Tomizaki, T.; Yamaguchi, H.; Shinzawa-ito, H.  
 submitted to the Brookhaven Protein Data Bank, April 1996  
 A:Reference number: A67451; PDB:1OCC  
 A:Contents: annotation; X-ray crystallography, 2.8 angstroms, residues 1-109  
 R:Tsuikihara, T.; Aoyama, H.; Yamashita, E.; Tomizaki, T.; Yamaguchi, H.; Shinzawa-Ittoh, H.  
 Science 272, 1136-1144, 1996  
 A:Title: The whole structure of the 13-subunit oxidized cytochrome c oxidase at 2.8 ang  
 A:Reference number: A57981; MUID:96216288; PMID:8638158  
 A:Contents: annotation; X-ray crystallography, 2.8 angstroms  
 C:Genetics:

A:Genome: nuclear  
 C:Complex: part of a 13 chain complex spanning the inner mitochondrial membrane and cons  
 Vb (see PIR:OGB06A), Via (see PIR:OGB06), VId (see PIR:OGB07), VIC (see PIR:OGB06C), V  
 dimers within the mitochondrial inner-membrane  
 C:Function:  
 A:Description: the cytochrome-c oxidase complex catalyzes the oxidation of four molecule  
 ns from the mitochondrial matrix producing two molecules of water and lowering the conce  
 A:Pathway: oxidative phosphorylation; respiratory chain  
 A:Note: the role of chain Va is not clear  
 C:Superfamily: mammalian cytochrome-c oxidase chain Va

C:Keywords: electron transfer; membrane-associated complex; mitochondrial inner membran  
 F:1-109/Product: cytochrome-c oxidase chain Va #status experimental <MAT>  
 F:1-109/Domain: mitochondrial matrix #status experimental <MM1>

Query Match 95.2%; Score 20; DB 1; Length 109;

Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
 ||||  
 Db 56 RRLN 59

## RESULT 36

## AF0301

conserved hypothetical protein YPO2469 [imported] - Yersinia pestis (strain CO92)  
 C:Species: Yersinia pestis  
 C>Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004  
 C:Accession: AF0301  
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.  
 deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
 il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,  
 Nature 413, 523-527, 2001  
 A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360  
 A:Accession: AF0301  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-110 <KUR>  
 A:Cross-references: UNIPROT:Q8ZDT7; UNIPARC:UPI00000CD916; GB:AL590842; PIDN:CAC91274.1  
 C:Genetics:  
 A:Gene: YPO2469

Query Match 95.2%; Score 20; DB 2; Length 110;

Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
 ||||  
 Db 4 RRLN 7

## RESULT 37

## F71149

hypothetical protein PH0408 - Pyrococcus horikoshii

C:Species: Pyrococcus horikoshii  
 C>Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 12-Jul-2004  
 C:Accession: F71149  
 R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Seki  
 M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguch  
 DNA Res. 5, 55-76, 1998

A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic  
 A:Reference number: A71000; MUID:98344137; PMID:9679194

A:Accession: F71149

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-112 <KAW>

A:Cross-references: UNIPROT:O58145; UNIPARC:UPI0000062DE3; GB:AP000002; NID:g3236129; P

A:Experimental source: strain O73

A:Note: this accession replaces an interim accession for a sequence replaced by GenBank

C:Genetics:

A:Gene: PH0408

Query Match 95.2%; Score 20; DB 2; Length 112;

Best Local Similarity 100.0%; Pred. No. 2.7e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
 ||||  
 Db 21 RRLN 24

## RESULT 38

E83073  
hypothetical protein PA4575 [imported] - *Pseudomonas aeruginosa* (strain PA01)  
C:Species: *Pseudomonas aeruginosa*  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
C:Accession: E83073  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bz  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho  
A:Reference number: AB2950; MUID:20437337; PMID:10984043  
A:Accession: E83073  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-113 <STO>  
A:Cross-references: UNIPROT:Q9HVK9; UNIPARC:UPI00000C5D6D; GB:AE0040871; GB:AE004091; NID  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA4575

Query Match 95.2%; Score 20; DB 2; Length 113;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
|||||  
DB 94 RRLN 97

RESULT 39  
AD2868  
hypothetical protein Atu2374 [imported] - *Agrobacterium tumefaciens* (strain C58, Dupont)  
C:Species: *Agrobacterium tumefaciens*  
C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004  
C:Accession: AD2868  
R:Wood, D.W.; Secubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell  
; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
ster, E.W.  
A:Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.  
A:Reference number: AB2577; MUID:21608550; PMID:11743193  
A:Accession: AD2868  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-113 <KUR>  
A:Cross-references: UNIPROT:Q8UCV9; UNIPARC:UPI00001646FB; GB:AE008688; PIDN:AAL43362.1;  
A:Experimental source: strain C58 (Dupont)  
C:Genetics:  
A:Gene: Atu2374  
A:Map position: circular chromosome

Query Match 95.2%; Score 20; DB 2; Length 113;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
|||||  
DB 99 RRLN 102

RESULT 40  
AI0294  
Probable membrane protein YPO2417 [imported] - *Yersinia pestis* (strain CO92)  
C:Species: *Yersinia pestis*  
C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004  
C:Accession: AI0294  
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.  
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,  
Nature 413, 523-527, 2001  
A:Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360  
A:Accession: AI0294  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-114 <KUR>  
A:Cross-references: UNIPROT:Q93PD5; UNIPARC:UPI00000D0133; GB:AL590842; PIDN:CAC91221.1;  
C:Genetics:  
A:Gene: YPO2417

Query Match 95.2%; Score 20; DB 2; Length 114;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
|||||  
DB 34 RRLN 37

RESULT 41  
AF0052  
hypothetical protein YPO0423 [imported] - *Yersinia pestis* (strain CO92)  
C:Species: *Yersinia pestis*  
C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004  
C:Accession: AF0052  
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B  
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,  
Nature 413, 523-527, 2001  
A:Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.  
A:Reference number: AB0001; MUID:21470413; PMID:11586360  
A:Accession: AF0052  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-115 <KUR>  
A:Cross-references: UNIPROT:Q8ZIR6; UNIPARC:UPI00000DC920; GB:AL590842; PIDN:CAC89281.1;  
C:Genetics:  
A:Gene: YPO0423

Query Match 95.2%; Score 20; DB 2; Length 115;  
Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
|||||  
DB 30 RRLN 33

RESULT 42  
S64490  
hypothetical protein YGR176w - yeast (*Saccharomyces cerevisiae*)  
N:Alternate names: hypothetical protein G7I31  
C:Species: *Saccharomyces cerevisiae*  
C:Date: 17-May-1996 #sequence\_revision 17-May-1996 #text\_change 09-Jul-2004  
C:Accession: S64490; PS0248  
R:Hebling, U.; Hofmann, B.; Delius, H.  
submitted to the Protein Sequence Database, May 1996  
A:Reference number: S64003  
A:Accession: S64490  
A:Molecule type: DNA  
A:Residues: 1-115 <HEB>  
A:Cross-references: UNIPROT:P32475; UNIPARC:UPI000013AF8D; EMBL:Z72961; NID:GI323311; P  
A:Experimental source: strain S288C  
R:Jandrositz, A.; Turnowsky, F.; Hoegenauer, G.  
Gene 107, 155-160, 1991  
A:Title: The gene encoding squalene epoxidase from *Saccharomyces cerevisiae*: cloning an  
A:Reference number: JH0490; MUID:92077427; PMID:1743514  
A:Accession: PS0248  
A:Molecule type: DNA  
A:Residues: 1-55; NSPR <JAN>  
A:Cross-references: UNIPARC:UPI000017CCAC; GB:M64994  
C:Genetics:  
A:Cross-references: SGD:S0003408  
A:Map position: 7R



Query Match 95.2%; Score 20; DB 2; Length 115;  
Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
Db 95 RRLN 98

RESULT 43  
T18073  
hypothetical protein A571R - Chlorella virus PBCV-1  
C:Species: Chlorella virus PBCV-1  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T18073  
R:Graves, M.V.; Van Etten, J.L.  
submitted to the EMBL Data Library, May 1999  
A:Reference number: Z18806  
A:Accession: T18073  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-116 <GRA>  
A:Cross-references: UNIPROT:O41053; UNIPARC:UPI00000F3C73; EMBL:U42580; NID:g4028896; PDB:1W33  
A:Experimental source: specific host Chlorella strain NC64A  
C:Genetics:  
A:Note: A571R

Query Match 95.2%; Score 20; DB 2; Length 116;  
Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
Db 96 RRLN 99

RESULT 44  
H97644  
hypothetical protein AGR\_C\_4309 [imported] - Agrobacterium tumefaciens (strain C58, Cere  
C:Species: Agrobacterium tumefaciens  
C>Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 09-Jul-2004  
C:Accession: H97644  
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,  
A.; Liu, F.; Wollan, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;  
Science 294, 2323-2328, 2001  
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum  
A:Reference number: A97359; MUID:21608551; PMID:11743194  
A:Accession: H97644  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-118 <KUR>  
A:Cross-references: UNIPROT:Q8UCV9; UNIPARC:UPI00000D1E9B; GB:AE007869; PDB:AAK88113.1;  
C:Genetics:  
A:Gene: AGR\_C\_4309  
A:Map position: circular chromosome

Query Match 95.2%; Score 20; DB 2; Length 118;  
Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
Db 104 RRLN 107

RESULT 45  
B72663  
hypothetical protein APE0730 - Aeropyrum pernix (strain K1)  
C:Species: Aeropyrum pernix  
C>Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
C:Accession: B72663  
R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;  
DNA Res. 6, 83-101, 1999  
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy  
A:Reference number: A72450; MUID:99310339; PMID:10382966  
A:Accession: B72663  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-121 <KAW>  
A:Cross-references: UNIPROT:Q9YE40; UNIPARC:UPI000005DC77; DDBJ:AP000060; NID:G5104188;  
A:Experimental source: strain K1  
C:Genetics:  
A:Gene: APE0730  
C:Superfamily: Aeropyrum pernix hypothetical protein APE0730

Query Match 95.2%; Score 20; DB 2; Length 121;  
Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
Db 67 RRLN 70

RESULT 46  
G70074  
hypothetical protein yxeE - Bacillus subtilis  
C:Species: Bacillus subtilis  
C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
C:Accession: G70074  
R:Kunat, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte  
C.; Bron, S.; Brouillet, S.; Bruchi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch  
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galle  
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F  
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois  
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauee  
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetell  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon  
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Sero  
akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpestra, P.; Toognoni, A.; Toato, V.; Uchiyama  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,  
A:Authors: Yoshikawa, H.F.; Zumschein, E.; Yoshikawa, H.; Danchin, A.  
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
A:Reference number: A69580; MUID:98044033; PMID:9384377  
A:Accession: G70074  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-121 <KUN>  
A:Cross-references: UNIPROT:P54944; UNIPARC:UPI0000060CB2; GB:Z99124; GB:AL009126; NID:  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: yxeE

Query Match 95.2%; Score 20; DB 2; Length 121;  
Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
Db 94 RRLN 97

RESULT 47  
G01477  
ribosomal protein L35 - human  
C:Species: Homo sapiens (man)  
C>Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 09-Jul-2004  
C:Accession: G01477  
R:Patel, S.K.  
submitted to the EMBL Data Library, July 1994  
A:Reference number: G07290  
A:Accession: G01477

A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-123 <PAT>  
A:Cross-references: UNIPROT:P42766; UNIPARC:UPI000015A4DD; EMBL:U12465; NID:G562073; PID  
C:Superfamily: rat ribosomal protein L35

Query Match 95.2%; Score 20; DB 2; Length 123;  
Best Local Similarity 100.0%; Pred. No. 3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
|||  
Db 93 RRLN 96

RESULT 48  
B95328  
hypothetical protein Sma0983 [imported] - Sinorhizobium meliloti (strain 1021) magaplas  
C:Species: Sinorhizobium meliloti  
C>Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 09-Jul-2004  
C:Accession: B95328  
R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bow  
; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.  
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001  
A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilot  
A:Reference number: A95262; MUID:21396509; PMID:11481432  
A:Accession: B95328  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-123 <KUR>  
A:Cross-references: UNIPROT:Q92Z86; UNIPARC:UPI00000CB109; GB:AE006469; PIDN:AAK65188.1;  
A:Experimental source: strain 1021, megaplasmid pSymA  
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, P.; Barloy-Hubler,  
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;  
L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001  
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,  
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wdhg, K.; Yeh, K.  
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.  
A:Reference number: A96039; MUID:21368234; PMID:11474104  
A:Contents: annotation  
C:Genetics:  
A:Gene: Sma0983  
A:Genome: plasmid

Query Match 95.2%; Score 20; DB 2; Length 123;  
Best Local Similarity 100.0%; Pred. No. 3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
|||  
Db 55 RRLN 58

RESULT 49  
T37118  
probable transposase, truncated [imported] - Streptomyces coelicolor  
C:Species: Streptomyces coelicolor  
C>Date: 08-Sep-2000 #sequence\_revision 08-Sep-2000 #text\_change 15-Sep-2000  
C:Accession: T37118  
R:Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.  
submitted to the EMBL Data Library, August 1999  
A:Reference number: Z21588  
A:Accession: T37118  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-124 <SAU>  
A:Cross-references: UNIPARC:UPI00001795PB; EMBL:AL109950; PIDN:CAB52969.1; GSPDB:GN00070  
A:Experimental source: strain A3(2)  
C:Genetics:  
A:Gene: SCOEDB:SCJ4.35  
C:Superfamily: Streptomyces coelicolor probable transposase SC6G9.36c

Query Match 95.2%; Score 20; DB 2; Length 124;  
Best Local Similarity 100.0%; Pred. No. 3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
|||  
Db 116 RRLN 119

RESULT 50  
AF2757  
conserved hypothetical protein Atui470 [imported] - Agrobacterium tumefaciens (strain C58)  
C:Species: Agrobacterium tumefaciens  
C>Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004  
C:Accession: AF2757  
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I  
; Karp, P.; Romero, P.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell  
erage, G.; Gillet, W.; Zhang, S.  
Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tsao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
ster, E.W.  
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A:Reference number: AB2577; MUID:21608550; PMID:11743193  
A:Accession: AF2757  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-125 <KUR>  
A:Cross-references: UNIPROT:Q8UFC8; UNIPARC:UPI00001283E7; GB:AE008688; PIDN:AAL42476.1  
A:Experimental source: strain C58 (Dupont)  
C:Genetics:  
A:Gene: Atui470  
A:Map position: circular chromosome  
C:Superfamily: hypothetical protein MJ1523

Query Match 95.2%; Score 20; DB 2; Length 125;  
Best Local Similarity 100.0%; Pred. No. 3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
|||  
Db 57 RRLN 60

Search completed: January 25, 2006, 18:41:53  
Job time : 38.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:33:23 ; Search time 63 Seconds  
(without alignments)  
55.994 Million cell updates/sec

Title: US-10-771-242-295  
Perfect score: 21  
Sequence: 1 RRLNX 5

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :

Uniprot\_05.80.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	95.2	16	2	Q23912 dictyosteli
2	20	95.2	27	2	Q5BX77 SCHJA
3	20	95.2	33	2	Q50L67 DROBP
4	20	95.2	33	2	Q50L79_9DIPT
5	20	95.2	33	2	Q50L83_9DIPT
6	20	95.2	33	2	Q50L86_9DIPT
7	20	95.2	33	2	Q50L89_9DIPT
8	20	95.2	34	2	Q50L77 DROAN
9	20	95.2	35	2	Q4ZG64 HUMAN
10	20	95.2	35	2	Q888U8_PSESM
11	20	95.2	35	2	Q4SVAL_TETNG
12	20	95.2	36	2	Q4Z5P5_PLABE
13	20	95.2	36	2	Q7ULM0_RHOBA
14	20	95.2	37	2	Q7SHZ1_NEUCR
15	20	95.2	37	2	Q8F419_LEPIN
16	20	95.2	39	2	Q4TTJ0_STRPU
17	20	95.2	39	2	Q4TTJ2_STRPU
18	20	95.2	39	2	Q4TTJ4_STRPU
19	20	95.2	39	2	Q4TTJ8_STRPU
20	20	95.2	39	2	Q4TTK1_STRPU
21	20	95.2	39	2	Q66546_9GAMA
22	20	95.2	40	2	Q4TTG6_STRDR
23	20	95.2	40	2	Q663L4_YERPS
24	20	95.2	43	2	Q7RL48_PLAYO
25	20	95.2	43	2	Q8KEJ5_CHLTE
26	20	95.2	44	2	Q5CSH2_SCHJA
27	20	95.2	44	2	Q72A18_DESVH
28	20	95.2	47	2	Q99IR3_9GEMI
29	20	95.2	47	2	Q99IR5_tomato leaf
30	20	95.2	48	2	Q99IR1_soybean cri
31	20	95.2	50	1	PR21_SCYCA

32	20	95.2	50	2	Q6DTN0_CANGA	O6dtn0 candida gla
33	20	95.2	50	2	Q83LH1_ENTFA	O4t3h1 enterococcu
34	20	95.2	50	2	Q479A7_TETNG	O4t9a7 tetraodon n
35	20	95.2	51	2	Q9ZSF6_9SOLA	O9zsf6 lycium ande
36	20	95.2	51	2	Q8RTK9_LEUME	O8rtk9 leuconostoc
37	20	95.2	51	2	Q8U4Z9_AGR5	O8u4z9 agrobacteri
38	20	95.2	51	2	Q57N53_SALCH	O57n53 salmonella
39	20	95.2	51	2	Q8VKJ1_MYCTU	O8vkj1 mycobacteri
40	20	95.2	51	2	Q5P112_SALPA	O5p112 salmonella
41	20	95.2	51	2	Q8Z5T2_SALTI	O8z5t2 salmonella
42	20	95.2	51	2	Q8ZNT9_SALTY	O8znt9 salmonella
43	20	95.2	52	2	Q96142_HUMAN	O96142 homo sapien
44	20	95.2	52	2	Q54BT9_DICDI	O54bt9 dictyosteli
45	20	95.2	53	2	Q9ZSG7_9SOLA	O9zsg7 lycium ande
46	20	95.2	54	2	Q8EDQ1_SHEON	O8edq1 shewanella
47	20	95.2	54	2	Q8DI57_SYNEL	O8di57 synecococc
48	20	95.2	54	2	Q9E3W1_TYLCV	O9e3w1 tomato yell
49	20	95.2	54	2	Q4GXV3_9GEMI	O4gxv3 tomato gemi
50	20	95.2	55	2	Q4GXV7_9GEMI	O4gxv7 tomato gemi
51	20	95.2	56	2	Q5JVT5_HUMAN	O5jvt5 homo sapien
52	20	95.2	56	2	Q4GXV1_9GEMI	O4gxv1 tomato gemi
53	20	95.2	58	2	Q6UU04_ORYSA	O6uu04 oryza sativ
54	20	95.2	58	2	Q6PD70_ACIAD	O6pd70 acinetobact
55	20	95.2	59	2	Q6JCV6_XANGC	O6jcv6 xanthomonas
56	20	95.2	59	2	Q4GXV9_9GEMI	O4gxv9 tomato gemi
57	20	95.2	60	2	Q64A06_9ARCH	O64a06 uncultured
58	20	95.2	60	2	Q4LVY7_9BURK	O4lvv7 burkholderi
59	20	95.2	60	2	Q8XV30_RALSO	O8xv30 ralstonia s
60	20	95.2	60	2	Q62GM3_BURMA	O62gm3 burkholderi
61	20	95.2	60	2	Q63Q29_BURPS	O63q29 burkholderi
62	20	95.2	60	2	Q68UX3_9GEMI	O68ux3 ageratum ye
63	20	95.2	61	2	Q8WYD0_HUMAN	O8wyd0 homo sapien
64	20	95.2	61	2	Q82X74_NITEU	O82x74 nitrosomona
65	20	95.2	61	2	Q98J39_RHILO	O98j39 rhizobium l
66	20	95.2	61	2	Q9K743_BACHD	O9k743 bacillus ha
67	20	95.2	62	2	Q5WHN0_BACSK	O5whn0 bacillus cl
68	20	95.2	62	2	Q4GXV5_9GEMI	O4gxv5 tomato gemi
69	20	95.2	62	2	Q5BSP2_SCHJA	O5bsp2 schistosoma
70	20	95.2	71	20	Q663M1_YERPS	O663m1 yersinia ps
71	20	95.2	63	2	Q9QCK2_BDV	O9qck2 borna disea
72	20	95.2	63	2	Q9QCK5_BDV	O9qck5 borna disea
73	20	95.2	63	2	Q9QCK8_BDV	O9qck8 borna disea
74	20	95.2	63	2	Q9QCL1_BDV	O9qcl1 borna disea
75	20	95.2	63	2	Q9QCL4_BDV	O9qcl4 borna disea
76	20	95.2	63	2	Q9QCL7_BDV	O9qcl7 borna disea
77	20	95.2	63	2	Q9QCM0_BDV	O9qcm0 borna disea
78	20	95.2	63	2	Q9QCM3_BDV	O9qcm3 borna disea
79	20	95.2	63	2	Q6ADC8_LEIXX	O6adc8 leifsonia x
80	20	95.2	64	2	Q6ZX71_9GEMI	O6zx71 papaya leaf
81	20	95.2	64	2	Q6ZX75_9GEMI	O6zx75 papaya leaf
82	20	95.2	64	2	Q6ZX79_9GEMI	O6zx79 tomato leaf
83	20	95.2	65	2	Q84KT4_BRAOL	O84kt4 brassica ol
84	20	95.2	65	2	Q84KT4_BRAOL	O84kt4 brassica ol
85	20	95.2	65	2	Q6ZX87_9GEMI	O6zx87 tomato leaf
86	20	95.2	65	2	Q6ZX89_9GEMI	O6zx89 ageratum ye
87	20	95.2	65	2	Q934F4_TREPH	O934f4 treponema p
88	20	95.2	66	2	Q93DL5_TREPA	O93dl5 treponema p
89	20	95.2	66	2	Q93DL6_9SPIO	O93dl6 spirochete
90	20	95.2	66	2	Q93DL7_TREMD	O93dl7 treponema m
91	20	95.2	66	2	Q93DL8_TRESO	O93dl8 treponema s
92	20	95.2	66	2	Q93DL9_TREDE	O93dl9 treponema d
93	20	95.2	66	2	Q93DM0_9SPIO	O93dm0 spirochete
94	20	95.2	66	2	Q93DM1_9SPIO	O93dm1 spirochete
95	20	95.2	66	2	Q93DM2_9SPIO	O93dm2 spirochete
96	20	95.2	66	2	Q8ABX4_BACTN	O8abx4 bacteroides
97	20	95.2	66	2	Q68UV4_9GEMI	O68uv4 tomato leaf
98	20	95.2	66	2	Q6ZX73_9GEMI	O6zx73 papaya leaf
99	20	95.2	66	2	Q6ZX85_9GEMI	O6zx85 tomato leaf
100	20	95.2	67	2	Q6NS35_HUMAN	O6ns35 homo sapien
101	20	95.2	67	2	Q5TM12_ANOGA	O5tm12 anopheles g
102	20	95.2	67	2	Q41381_9ASTRI	O41381 senecio odo
103	20	95.2	67	2	Q6ZX83_9GEMI	O6zx83 tomato leaf
104	20	95.2	67	2	Q6ZX83_9GEMI	O6zx83 tomato leaf

105	20	95.2	68	2	Q757Q8	ashbya_gosa	Q757q8	178	20	95.2	74	2	Q8B8P9	9GEMI	Q8b8p9	tomato gemi
106	20	95.2	68	2	Q4G0U0	homo sapien	Q4g0u0	179	20	95.2	74	2	Q8B8Q0	9GEMI	Q8b8q0	tomato gemi
107	20	95.2	68	2	Q7X6G5	oryza sativ	Q7x6g5	180	20	95.2	74	2	Q8B8Q1	9GEMI	Q8b8q1	tomato gemi
108	20	95.2	68	2	Q7F9H3	oryza sativ	Q7f9h3	181	20	95.2	74	2	Q8B9F9	NPVRO	Q8b9f9	rachiplusia
109	20	95.2	68	2	Q6HFQ9	bacillus th	Q6hfq9	182	20	95.2	74	2	Q8UYP5	9GEMI	Q8uyp5	tobacco lea
110	20	95.2	68	2	Q926Q2	tylcv	Q926q2	183	20	95.2	74	2	Q8UYP7	9GEMI	Q8uyp7	tobacco lea
111	20	95.2	68	2	Q6KCL1	9GEMI	Q6kcl1	184	20	95.2	74	2	Q4VYJ1	9GEMI	Q4vyj1	tomato yell
112	20	95.2	68	2	Q6KCL3	9GEMI	Q6kcl3	185	20	95.2	74	2	Q4VYJ3	9GEMI	Q4vyj3	tomato yell
113	20	95.2	68	2	Q6KCL5	9GEMI	Q6kcl5	186	20	95.2	74	2	Q4VYJ5	9GEMI	Q4vyj5	tomato yell
114	20	95.2	68	2	Q6ZX59	9GEMI	Q6zx59	187	20	95.2	74	2	Q4VYJ7	9GEMI	Q4vyj7	tomato yell
115	20	95.2	68	2	Q6ZX61	9GEMI	Q6zx61	188	20	95.2	74	2	Q4VYJ9	9GEMI	Q4vyj9	tomato yell
116	20	95.2	68	2	Q6ZX63	9GEMI	Q6zx63	189	20	95.2	74	2	Q4VYK1	9GEMI	Q4vyk1	tomato yell
117	20	95.2	68	2	Q6ZX67	9GEMI	Q6zx67	190	20	95.2	74	2	Q4VYK3	9GEMI	Q4vyk3	tomato yell
118	20	95.2	68	2	Q6ZX69	9GEMI	Q6zx69	191	20	95.2	74	2	Q4VYK5	9GEMI	Q4vyk5	tomato yell
119	20	95.2	68	2	Q6ZX77	9GEMI	Q6zx77	192	20	95.2	74	2	Q4VYK7	9GEMI	Q4vyk7	tomato yell
120	20	95.2	68	2	Q6ZX81	9GEMI	Q6zx81	193	20	95.2	74	2	Q4VYK9	9GEMI	Q4vyk9	tomato yell
121	20	95.2	68	2	Q8UZB8	9GEMI	Q8uzb8	194	20	95.2	74	2	Q4VYL1	9GEMI	Q4vy11	tobacco cur
122	20	95.2	68	2	Q8UZC0	9GEMI	Q8uzc0	195	20	95.2	74	2	Q4VYL3	9GEMI	Q4vy13	tobacco cur
123	20	95.2	68	2	Q9YKD4	TYLCV	Q9ykd4	196	20	95.2	74	2	Q4VYL7	9GEMI	Q4vy17	tobacco cur
124	20	95.2	68	2	Q6BGY5	DEBHA	Q6bg55	197	20	95.2	74	2	Q4VYL9	9GEMI	Q4vy19	tobacco lea
125	20	95.2	69	2	Q6ZX55	9GEMI	Q6zx55	198	20	95.2	74	2	Q4VYV1	9GEMI	Q4vym1	tobacco lea
126	20	95.2	69	2	Q6ZX65	9GEMI	Q6zx65	199	20	95.2	74	2	Q4VYV3	9GEMI	Q4vym3	tobacco lea
127	20	95.2	70	2	Q7WHS6	BORBR	Q7whs6	200	20	95.2	74	2	Q4VYV5	9GEMI	Q4vym5	tobacco lea
128	20	95.2	70	2	Q5TJC8	9GEMI	Q5tjc8	201	20	95.2	74	2	Q4VYV7	9GEMI	Q4vym7	tobacco lea
129	20	95.2	70	2	Q5TJD0	9GEMI	Q5tjd0	202	20	95.2	74	2	Q4VYV9	9GEMI	Q4vym9	malvastrum
130	20	95.2	70	2	Q6ZX57	9GEMI	Q6zx57	203	20	95.2	74	2	Q4VYV1	9GEMI	Q4vyn1	malvastrum
131	20	95.2	70	2	Q8B8P3	9GEMI	Q8b8p3	204	20	95.2	74	2	Q4VYV3	9GEMI	Q4vyn3	malvastrum
132	20	95.2	70	2	Q8B8P4	9GEMI	Q8b8p4	205	20	95.2	74	2	Q4VYV5	9GEMI	Q4vyn5	malvastrum
133	20	95.2	70	2	Q8B8P6	9GEMI	Q8b8p6	206	20	95.2	74	2	Q4VYV7	9GEMI	Q4vyn7	malvastrum
134	20	95.2	70	2	Q8B8P7	9GEMI	Q8b8p7	207	20	95.2	74	2	Q4VYV1	9GEMI	Q4vyp1	malvastrum
135	20	95.2	70	2	Q8B8P8	9GEMI	Q8b8p8	208	20	95.2	74	2	Q4VYV3	9GEMI	Q4vyp3	malvastrum
136	20	95.2	70	2	Q8B8P8	9GEMI	Q8b8p8	209	20	95.2	74	2	Q4VYV5	9GEMI	Q4vyp5	malvastrum
137	20	95.2	71	1	TRAD RHISN		Q4stl0	210	20	95.2	74	2	Q4VYP7	9GEMI	Q4vyp7	malvastrum
138	20	95.2	71	1	Y1435	METJA	Q58830	211	20	95.2	74	2	Q4VYP9	9GEMI	Q4vyp9	malvastrum
139	20	95.2	71	2	Q8G9P7	ENTAG	Q8g9p7	212	20	95.2	74	2	Q4VYQ1	9GEMI	Q4vyq1	malvastrum
140	20	95.2	71	2	Q5JZL7	9GEMI	Q5jlz7	213	20	95.2	74	2	Q4VYQ5	9GEMI	Q4vyq5	tomato yell
141	20	95.2	71	2	Q68U08	9GEMI	Q68u08	214	20	95.2	74	2	Q4VYQ7	9GEMI	Q4vyq7	tomato yell
142	20	95.2	72	2	Q7MBK8	VIBVY	Q7mbk8	215	20	95.2	74	2	Q4VYQ9	9GEMI	Q4vyq9	tomato yell
143	20	95.2	72	2	Q9R120	YERPE	Q9r120	216	20	95.2	74	2	Q4VYR3	9GEMI	Q4vyr3	tomato yell
144	20	95.2	72	2	Q5TJD2	9GEMI	Q5tjd2	217	20	95.2	74	2	Q4VYR5	9GEMI	Q4vyr5	tomato yell
145	20	95.2	72	2	Q6R2T8	PASDO	Q6r2t8	218	20	95.2	74	2	Q4GXW1	9GEMI	Q4gxw1	tomato gemi
146	20	95.2	72	2	Q7Y2A6	9CAUD	Q7y2a6	219	20	95.2	75	2	Q6Z8U5	ORYSA	Q6z8u5	oryza sativ
147	20	95.2	73	2	Q7Y2N3	9CAUD	Q7y2n3	220	20	95.2	75	2	Q9S0S0	STRCO	Q9s0s0	streptococ
148	20	95.2	73	2	Q7Y344	9CAUD	Q7y344	221	20	95.2	75	2	Q9A0Q9	STRPY	Q9a0q9	streptococ
149	20	95.2	73	2	Q8SC56	9CAUD	Q8sc56	222	20	95.2	75	2	Q5WM52	9GEMI	Q5wm52	papaya leaf
150	20	95.2	73	2	Q8X3F3	EC057	Q8x3f3	223	20	95.2	75	2	Q5WM54	9GEMI	Q5wm54	papaya leaf
151	20	95.2	73	2	Q9EYAL	EC057	Q9eyal	224	20	95.2	75	2	Q5WM70	9GEMI	Q5wm70	ageratum ye
152	20	95.2	73	2	Q9KXG7	EC057	Q9kxg7	225	20	95.2	75	2	Q5WM72	9GEMI	Q5wm72	ageratum ye
153	20	95.2	73	2	Q7TBL8	9GEMI	Q7tbl8	226	20	95.2	75	2	Q5WM74	9GEMI	Q5wm74	ageratum ye
154	20	95.2	73	2	Q4Z1J7	9PARA	Q4z1j7	227	20	95.2	75	2	Q5WM76	9GEMI	Q5wm76	ageratum ye
155	20	95.2	73	2	Q4Z1J8	9PARA	Q4z1j8	228	20	95.2	75	2	Q5WM78	9GEMI	Q5wm78	ageratum ye
156	20	95.2	73	2	Q4Z1J9	9PARA	Q4z1j9	229	20	95.2	75	2	Q5WM80	9GEMI	Q5wm80	ageratum ye
157	20	95.2	73	2	Q4Z1K0	9PARA	Q4z1k0	230	20	95.2	75	2	Q5WM82	9GEMI	Q5wm82	ageratum ye
158	20	95.2	73	2	Q4Z1K1	9PARA	Q4z1k1	231	20	95.2	75	2	Q5WM84	9GEMI	Q5wm84	ageratum ye
159	20	95.2	74	2	Q8R8F5	THETN	Q8r8f5	232	20	95.2	75	2	Q5WM86	9GEMI	Q5wm86	ageratum ye
160	20	95.2	74	2	Q9PGQ6	XYLFA	Q9pgq6	233	20	95.2	75	2	Q5WM88	9GEMI	Q5wm88	ageratum ye
161	20	95.2	74	2	Q68UV2	9GEMI	Q68uv2	234	20	95.2	75	2	Q5WM90	9GEMI	Q5wm90	ageratum ye
162	20	95.2	74	2	Q7TGS1	9GEMI	Q7tgs1	235	20	95.2	75	2	Q5WM92	9GEMI	Q5wm92	ageratum ye
163	20	95.2	74	2	Q8B6S8	TYLCV	Q8b6s8	236	20	95.2	75	2	Q5WM94	9GEMI	Q5wm94	ageratum ye
164	20	95.2	74	2	Q8B6T0	TYLCV	Q8b6t0	237	20	95.2	75	2	Q5WM98	9GEMI	Q5wm98	ageratum ye
165	20	95.2	74	2	Q8B6T2	TYLCV	Q8b6t2	238	20	95.2	75	2	Q5WMA0	9GEMI	Q5wma0	ageratum ye
166	20	95.2	74	2	Q8B6V7	9GEMI	Q8b6v7	239	20	95.2	75	2	Q6A0F5	9GEMI	Q6a0f5	stachytarph
167	20	95.2	74	2	Q8B6V8	9GEMI	Q8b6v8	240	20	95.2	75	2	Q6A0F7	9GEMI	Q6a0f7	stachytarph
168	20	95.2	74	2	Q8B6W0	9GEMI	Q8b6w0	241	20	95.2	75	2	Q6A0G1	9GEMI	Q6a0g1	stachytarph
169	20	95.2	74	2	Q8B6W1	9GEMI	Q8b6w1	242	20	95.2	75	2	Q6A0G3	9GEMI	Q6a0g3	stachytarph
170	20	95.2	74	2	Q8B6W2	9GEMI	Q8b6w2	243	20	95.2	75	2	Q6A0G5	9GEMI	Q6a0g5	stachytarph
171	20	95.2	74	2	Q8B6W4	9GEMI	Q8b6w4	244	20	95.2	75	2	Q6A0G7	9GEMI	Q6a0g7	stachytarph
172	20	95.2	74	2	Q8B6W6	9GEMI	Q8b6w6	245	20	95.2	75	2	Q6A0G9	9GEMI	Q6a0g9	stachytarph
173	20	95.2	74	2	Q8B6W7	9GEMI	Q8b6w7	246	20	95.2	75	2	Q7TGS2	9GEMI	Q7tgs2	malvastrum
174	20	95.2	74	2	Q8B6W8	9GEMI	Q8b6w8	247	20	95.2	75	2	Q8B8R1	9GEMI	Q8b8r1	ageratum ye
175	20	95.2	74	2	Q8B6X0	9GEMI	Q8b6x0	248	20	95.2	75	2	Q8B8R3	9GEMI	Q8b8r3	ageratum ye
176	20	95.2	74	2	Q8B6X1	9GEMI	Q8b6x1	249	20	95.2	75	2	Q9VRC9	9PARA	Q9vrc9	newcastle d
177	20	95.2	74	2	Q8B6X2	9GEMI	Q8b6x2	250	20	95.2	75	2				

251	20	95.2	75	2	Q68V32_9GEMI	Q68v32 sida yellow	324	20	95.2	81	1	OADG2_SALTY	P58650 salmonella
252	20	95.2	75	2	Q68V43_9GEMI	Q68v43 sida yellow	325	20	95.2	82	2	Q8GMW6_RAT	Q8gmw6 arabisopsis
253	20	95.2	75	2	Q4VYL5_9GEMI	Q4vyl5 tobacco cur	326	20	95.2	82	2	Q4KM25_RAT	Q4km25 rattus norv
254	20	95.2	75	2	Q4VZ16_9GEMI	Q4vz16 papaya leaf	327	20	95.2	83	2	Q54SS4_PLADI	Q54ss4 dictyosteli
255	20	95.2	75	2	Q4VZ18_9GEMI	Q4vz18 papaya leaf	328	20	95.2	83	2	Q4X4S6_PLADI	Q4x4s6 plasmodium
256	20	95.2	75	2	Q4VZ20_9GEMI	Q4vz20 ageratum ye	329	20	95.2	83	2	Q6K6S3_ORYSA	Q6k6s3 oryza sativ
257	20	95.2	75	2	Q4VZ23_9GEMI	Q4vz23 ageratum ye	330	20	95.2	83	2	Q6MXB9_SERMA	Q6mxb9 serratia ma
258	20	95.2	75	2	Q4U0F5_9GEMI	Q4u0f5 hibiscus ro	331	20	95.2	83	2	Q847N6_ASTYP	Q847n6 aster yello
259	20	95.2	75	2	Q4U0F7_9GEMI	Q4u0f7 lindernia a	332	20	95.2	83	2	Q4IVY5_AZOVI	Q4ivy5 azotobacter
260	20	95.2	75	2	Q4U0F9_9GEMI	Q4u0f9 lindernia a	333	20	95.2	83	2	Q7W6U0_BORPA	Q7w6u0 bordetella
261	20	95.2	75	2	Q5JZK9_9GEMI	Q5jzk9 senecio yel	334	20	95.2	85	2	Q9N114_SHEEP	Q9n114 ovis aries
262	20	95.2	75	2	Q5JZK7_9GEMI	Q5jzk7 senecio yel	335	20	95.2	85	2	Q96352_BRANA	Q96352 brassica na
263	20	95.2	76	2	Q79EL6_ECOLI	Q79el6 escherichia	336	20	95.2	85	2	Q5MOC8_STRT1	Q5moc8 streptococc
264	20	95.2	76	2	P76024_ECOLI	P76024 escherichia	337	20	95.2	85	2	Q5M4X5_STRT2	Q5m4x5 streptococc
266	20	95.2	76	2	Q5WM96_9GEMI	Q5wm96 ageratum ye	338	20	95.2	86	2	Q570C7_ATHAT	Q570c7 arabisopsis
267	20	95.2	77	2	Q64DT2_9ARCH	Q64dt2 uncultured	339	20	95.2	86	2	Q4LXV7_9BURK	Q4lxv7 burkholderi
268	20	95.2	77	2	Q5TUC9_9ANO	Q5tuc9 anopheles g	340	20	95.2	86	2	Q82UX2_NITEU	Q82ux2 nitrosomona
269	20	95.2	77	2	Q5OSR4_ENTHI	Q5osr4 entamoeba h	341	20	95.2	86	2	Q7UVES_RHOBA	Q7uve5 rhodopirell
270	20	95.2	77	2	Q36988_9PARA	Q36988 newcastle d	342	20	95.2	87	1	S8S1_SCVCA	S13275 scyllorhinu
271	20	95.2	77	2	Q36990_9PARA	Q36990 newcastle d	343	20	95.2	87	1	Q5CKC6_CRXHO	Q5ckc6 cryptospori
272	20	95.2	77	2	Q36992_9PARA	Q36992 newcastle d	344	20	95.2	87	2	Q94MP0_9CAUD	Q94mp0 bacterioph
273	20	95.2	77	2	Q36994_9PARA	Q36994 newcastle d	345	20	95.2	87	2	Q5GL76_BDV	Q5gl76 borna disea
274	20	95.2	77	2	Q36996_9PARA	Q36996 newcastle d	346	20	95.2	87	2	Q5GL97_BDV	Q5gl97 borna disea
275	20	95.2	77	2	Q36998_9PARA	Q36998 newcastle d	347	20	95.2	87	2	Q5GLA6_BDV	Q5glA6 borna disea
276	20	95.2	77	2	Q37000_9PARA	Q37000 newcastle d	348	20	95.2	87	2	Q5GLD0_BDV	Q5glD0 borna disea
277	20	95.2	77	2	Q37002_9PARA	Q37002 newcastle d	349	20	95.2	87	2	Q86622_BDV	Q86622 borna disea
278	20	95.2	77	2	Q37004_9PARA	Q37004 newcastle d	350	20	95.2	87	2	Q8JJK0_BDV	Q8jjk0 borna disea
279	20	95.2	77	2	Q37006_9PARA	Q37006 newcastle d	351	20	95.2	87	2	Q91222_BDV	Q91222 borna disea
280	20	95.2	77	2	Q37008_9PARA	Q37008 newcastle d	352	20	95.2	87	2	Q91225_BDV	Q91225 borna disea
281	20	95.2	77	2	Q37010_9PARA	Q37010 newcastle d	353	20	95.2	87	2	Q91229_BDV	Q91229 borna disea
282	20	95.2	77	2	Q37012_9PARA	Q37012 newcastle d	354	20	95.2	87	2	Q9Q9V0_BDV	Q9q9v0 borna disea
283	20	95.2	77	2	Q37014_9PARA	Q37014 newcastle d	355	20	95.2	87	2	Q9WNA1_BDV	Q9wna1 borna disea
284	20	95.2	77	2	Q68UY3_9GEMI	Q68uy3 papaya leaf	356	20	95.2	88	1	Y778_STRAS	Y8805 streptococc
285	20	95.2	77	2	Q6SEGE_9PARA	Q6sege newcastle d	357	20	95.2	88	1	Y798_STRAS	Y8829 streptococc
286	20	95.2	77	2	Q6SEG7_9PARA	Q6seg7 newcastle d	358	20	95.2	89	1	VP54_BPAPS	VP544 bacterioph
287	20	95.2	77	2	Q6SEG8_9PARA	Q6seg8 newcastle d	359	20	95.2	89	2	Q6H0V4_9CREN	Q6h0v4 sulfolobus
288	20	95.2	77	2	Q6SEG9_9PARA	Q6seg9 newcastle d	360	20	95.2	89	2	Q97WJ2_SULSO	Q97wj2 sulfolobus
289	20	95.2	77	2	Q83780_9PARA	Q83780 newcastle d	361	20	95.2	89	2	Q67TCS_SYMTH	Q67tc5 symbiobacte
290	20	95.2	77	2	Q83782_9PARA	Q83782 newcastle d	362	20	95.2	90	1	FETP_PSEAE	F90u36 pseudomonas
291	20	95.2	77	2	Q83790_9PARA	Q83790 newcastle d	363	20	95.2	90	1	FETP_PSESM	Q87uf5 pseudomonas
292	20	95.2	77	2	Q83800_9PARA	Q83800 newcastle d	364	20	95.2	90	2	Q5W667_ORYSA	Q5w667 oryza sativ
293	20	95.2	77	2	Q83820_9PARA	Q83820 newcastle d	365	20	95.2	90	2	Q4ZLP3_PSESY	Q4zlp3 pseudomonas
294	20	95.2	77	2	Q83824_9PARA	Q83824 newcastle d	366	20	95.2	90	2	Q4JZ28_AZOVI	Q4jz28 azotobacter
295	20	95.2	77	2	Q9DKI4_9PARA	Q9dk14 newcastle d	367	20	95.2	91	2	Q5TXS7_ANOVA	Q5txs7 anopheles g
296	20	95.2	77	2	Q9DKI6_9PARA	Q9dk16 newcastle d	368	20	95.2	91	2	Q6LW69_PHOPR	Q6lW69 photobacter
297	20	95.2	77	2	Q9DKI7_9PARA	Q9dk17 newcastle d	369	20	95.2	91	2	Q8D665_VIBUO	Q8d665 vibrio vuln
298	20	95.2	77	2	Q9DKI8_9PARA	Q9dk18 newcastle d	370	20	95.2	91	2	Q8R9Q8_THRTN	Q8r9q8 thermoanaer
299	20	95.2	77	2	Q9DKI9_9PARA	Q9dk19 newcastle d	371	20	95.2	91	2	Q92XL9_RHIME	Q92xl9 rhizobium m
300	20	95.2	77	2	Q9DKJ0_9PARA	Q9dkj0 newcastle d	372	20	95.2	91	2	Q7UM06_RHOBA	Q7um06 rhodopirell
301	20	95.2	77	2	Q9J3V8_9PARA	Q9j3v8 newcastle d	373	20	95.2	91	2	Q53CW8_9GAMA	Q53cw8 macaca fusc
302	20	95.2	77	2	Q9J3W0_9PARA	Q9j3w0 newcastle d	374	20	95.2	92	2	Q8ZB80_YERPE	Q8zb80 yersinia pe
303	20	95.2	77	2	Q9J3W4_9PARA	Q9j3w4 newcastle d	375	20	95.2	92	2	Q66FA3_YERPS	Q66fa3 yersinia ps
304	20	95.2	77	2	Q9J3X0_9PARA	Q9j3x0 newcastle d	376	20	95.2	92	2	Q8GZG5_STRAW	Q8gzg5 streptomyce
305	20	95.2	77	2	Q9J3X4_9PARA	Q9j3x4 newcastle d	377	20	95.2	92	2	Q8KBJ4_CHLITE	Q8kbj4 chlorobium
306	20	95.2	77	2	Q9J3X6_9PARA	Q9j3x6 newcastle d	378	20	95.2	93	2	Q5TYE6_ANOVA	Q5tye6 anopheles g
307	20	95.2	77	2	Q9J3X8_9PARA	Q9j3x8 newcastle d	379	20	95.2	93	2	Q9BLC2_CAEEL	Q9blc2 caenorhabdi
308	20	95.2	77	2	Q9J3Y0_9PARA	Q9j3y0 newcastle d	380	20	95.2	93	2	Q4QYB7_9ACTO	Q4qyb7 streptomyce
309	20	95.2	77	2	Q9QPC7_TYLCV	Q9qpc7 tomato yell	381	20	95.2	93	2	Q4QYC4_STRLI	Q4qyc4 streptomyce
310	20	95.2	78	2	Q7P2N2_SULTO	Q7p2n2 sulfolobus	382	20	95.2	93	2	Q4QYC5_9ACTO	Q4qyc5 streptomyce
311	20	95.2	78	2	Q7PL20_9ANO	Q7pl20 anopheles g	383	20	95.2	93	2	Q4QYC6_STRAM	Q4qyc6 streptomyce
312	20	95.2	78	2	Q5UKX2_9BROM	Q5ukx2 fragaria ch	384	20	95.2	94	1	ABCI_BPP22	Abci190 bacterioph
313	20	95.2	79	2	Q14891_HUMAN	Q14891 homo sapien	385	20	95.2	94	2	Q9VQ55_DROME	Q9vq55 drosophila
314	20	95.2	79	2	Q6P2M4_HUMAN	Q6p2m4 homo sapien	386	20	95.2	94	2	Q76H41_9CAUD	Q76h41 salmonella
315	20	95.2	79	2	Q4VYN9_9GEMI	Q4vyn9 malvastrum	387	20	95.2	94	2	Q8SCE8_9VIRU	Q8sce8 vibrio harv
316	20	95.2	79	2	Q4VTQ3_9GEMI	Q4vvtq3 malvastrum	388	20	95.2	94	2	Q8HAH6_BPST6	Q8hah6 bacterioph
317	20	95.2	79	2	Q8WPL8_9GEMI	Q8wpl8 tomato leaf	389	20	95.2	94	2	Q77D70_BP222	Q77d70 bacterioph
318	20	95.2	80	2	Q8WVT5_NARPS	Q8wvt5 narcissus p	390	20	95.2	94	2	Q5PF59_SALPA	Q5pf59 salmonella
319	20	95.2	80	2	Q57TK5_SALCH	Q57tk5 salmonella	391	20	95.2	94	2	Q8E1D3_STRAS	Q8e1d3 streptococc
320	20	95.2	80	2	Q8X2N0_EC057	Q8x2n0 escherichia	392	20	95.2	94	2	Q8B6V0_STRAS	Q8b6v0 streptococc
321	20	95.2	80	2	Q5WGQ0_BACSK	Q5wgq0 bacillus cl	393	20	95.2	95	2	Q7WZT3_BORGA	Q7wzt3 borrelia ga
322	20	95.2	80	2	Q82UA4_NITEU	Q82ua4 nitrosomona	394	20	95.2	95	2	Q6LHW7_PHOPR	Q6lhw7 photobacter
323	20	95.2	80	2	Q839M8_ENTPA	Q839m8 enterococcu	395	20	95.2	95	2	Q8ZF39_YERPE	Q8zf39 yersinia ps
							396	20	95.2	95	2	Q66B78_YERPS	Q66b78 yersinia ps

397	20	95.2	95	2	Q62B28_BURMA	Q62b28	burkholderi	470	20	95.2	108	2	Q88U03_LACPL	Q88u03	lactobacill
398	20	95.2	96	1	EXOX_RHISN	Fl185	rhizobium s	471	20	95.2	108	2	Q60516_9MURI	Q60516	rattus sp.
399	20	95.2	96	1	PAPA5_CARPA	P05993	carica papa	472	20	95.2	108	2	Q484B7_TYTING	Q484b7	tetraodon n
400	20	95.2	96	2	Q50ZD8_ENTHI	Q50zd8	entamoeba h	473	20	95.2	109	2	Q50UB4_ENTHI	Q50ub4	entamoeba n
401	20	95.2	96	2	Q8W6V6_9CAUD	Q8w6v6	cyanophaga	474	20	95.2	109	2	Q9ETL9_BORGA	Q9etl9	borrelia ga
402	20	95.2	96	2	Q7WZT4_BORAF	Q7wzt4	borrelia af	475	20	95.2	109	2	Q9ETM0_BORAF	Q9etm0	borrelia af
403	20	95.2	96	2	Q6W2H5_RHISN	Q6w2h5	rhizobium s	476	20	95.2	109	2	Q9ETM1_BORBU	Q9etm1	borrelia bu
404	20	95.2	96	2	Q5FNW8_GLUOX	Q5fnw8	gluconobact	477	20	95.2	109	2	Q9F9V8_BORBU	Q9f9v8	borrelia tu
405	20	95.2	96	2	Q6AJV7_DESPS	Q6ajv7	desulfotale	478	20	95.2	109	2	Q9F9V9_BORHE	Q9f9v9	borrelia he
406	20	95.2	97	2	Q6LWA7_PHOPR	Q6lwa7	photobacter	479	20	95.2	109	2	Q9F9W0_BORAD	Q9f9w0	borrelia an
407	20	95.2	97	2	Q8D647_VIBVU	Q8d647	vibrio vuln	480	20	95.2	109	2	Q9F9W1_9SP10	Q9f9w1	borrelia bi
408	20	95.2	97	2	EXOX_RHIME	Q24074	rhizobium m	481	20	95.2	109	2	Q9F9W2_9SP10	Q9f9w2	borrelia lu
409	20	95.2	98	1	Q9VUS9_DROME	Q9vus9	drosophila	482	20	95.2	109	2	Q9F9W3_BORJA	Q9f9w3	borrelia ja
410	20	95.2	98	2	Q5OR98_ENTHI	Q5or98	entamoeba h	483	20	95.2	109	2	Q9F9W4_9SP10	Q9f9w4	borrelia va
411	20	95.2	98	2	Q6EPF9_ORYSA	Q6epf9	oryza sativ	484	20	95.2	109	2	Q5X1S4_RAT	Q5x1s4	rattus norv
412	20	95.2	98	2	Q8EPC8_OCEIH	Q8epc8	oceanobacil	485	20	95.2	110	2	Q26682_METTH	Q26682	methanobact
413	20	95.2	98	2	Q8EPC8_OCEIH	Q8epc8	oceanobacil	486	20	95.2	110	2	Q59NH4_CANAL	Q59nh4	candida alb
414	20	95.2	98	2	Q8EPC8_OCEIH	Q8epc8	oceanobacil	487	20	95.2	110	2	Q9C1L6_NPUCR	Q9c1l6	neurospora
415	20	95.2	98	2	Q8EPC8_OCEIH	Q8epc8	oceanobacil	488	20	95.2	110	2	Q6YK33_BRANA	Q6y3k3	brassicica na
416	20	95.2	99	2	Q4Y8V2_PLACH	Q4y8v2	plasmodium	489	20	95.2	110	2	Q6YK33_BRANA	Q6y3k3	streptomyce
417	20	95.2	99	2	Q9L5I5_SALTI	Q9l5i5	salmonella	490	20	95.2	110	2	Q8YK33_RALSO	Q8y3k3	raistonia s
418	20	95.2	99	2	Q9Q0B0_NPVAG	Q9q0b0	anticarsia	491	20	95.2	110	2	Q8ZD77_YERPE	Q8zdt7	yersinia pe
419	20	95.2	99	2	Q9QMX1_9VIRU	Q9qmx1	hantaan vir	492	20	95.2	110	2	Q8C5D2_MOUSE	Q8c5d2	mus musculu
420	20	95.2	100	1	URE3_PROMM	Q7v3v4	prochloroco	493	20	95.2	110	2	Q5ZP61_9VIRU	Q5z6p1	cotesia con
421	20	95.2	100	1	URE3_PROMP	Q7v1b4	prochloroco	494	20	95.2	111	2	Q6NSH7_HUMAN	Q6nsh7	homo sapien
422	20	95.2	100	1	URE3_PROS9	Q7v1b4	prochloroco	495	20	95.2	111	2	Q5ZB16_ORYSA	Q5z6b16	oryza sativ
423	20	95.2	100	1	URE3_SYNVP	Q87400	synchococc	496	20	95.2	111	2	Q5ZDM6_ORYSA	Q5z6m6	oryza sativ
424	20	95.2	100	1	URE3_SYNVP	Q87400	synchococc	497	20	95.2	111	2	Q7WY00_PSEAE	Q7wy00	pseudomonas
425	20	95.2	100	2	Q5AXW8_EMENI	Q5axw8	aspergillus	498	20	95.2	111	2	Q5HLT5_STAEQ	Q5hlt5	staphylococ
426	20	95.2	100	2	Q27316_DROME	Q27316	drosophila	499	20	95.2	111	2	Q89CV9_BRAJA	Q89cv9	bradyrhizob
427	20	95.2	100	2	Q9VUS3_DROME	Q9vus3	drosophila	500	20	95.2	111	2	Q877N3_PSEPK	Q877n3	pseudomonas
428	20	95.2	100	2	Q5IGZ4_9CAUD	Q5igz4	bacterioph	501	20	95.2	111	2	Q877M6_PSEPK	Q877m6	pseudomonas
429	20	95.2	101	2	Q8I7Q4_DROME	Q8i7q4	drosophila	502	20	95.2	112	2	Q58145_PYRHO	Q58145	pyrococcus
430	20	95.2	101	2	Q8RZU9_ORYSA	Q8rzu9	oryza sativ	503	20	95.2	112	2	Q6BMJ2_DBBHA	Q6bmj2	debariomyce
431	20	95.2	101	2	Q6WKW5_9ACTO	Q6wk5	streptomyce	504	20	95.2	112	2	Q7QU11_GIALA	Q7qu11	giardia lam
432	20	95.2	101	2	Q6WL11_9ACTO	Q6wl11	streptomyce	505	20	95.2	112	2	Q7PUZ5_ANOGA	Q7puz5	anopheles g
433	20	95.2	101	2	Q6WL18_9ACTO	Q6wl18	streptomyce	506	20	95.2	112	2	Q69TP2_ORYSA	Q69tp2	oryza sativ
434	20	95.2	101	2	Q6WL38_STRST	Q6wl38	streptomyce	507	20	95.2	112	2	Q69TP2_ORYSA	Q69tp2	oryza sativ
435	20	95.2	101	2	Q6WL44_STRPH	Q6wl44	streptomyce	508	20	95.2	112	2	Q9FIG4_ARYTH	Q9fig4	arabidopsis
436	20	95.2	101	2	Q6WL62_STRCO	Q6wl62	streptomyce	509	20	95.2	112	2	Q84RD1_ARYTH	Q84rd1	arabidopsis
437	20	95.2	101	2	Q6WL69_STRAV	Q6wl69	streptomyce	510	20	95.2	112	2	Q4V2X5_BURMA	Q4v2x5	burkholderi
438	20	95.2	101	2	Q6WL73_STRAC	Q6wl73	streptomyce	511	20	95.2	112	2	Q91BD0_NPVST	Q91bd0	spodoptera
439	20	95.2	101	2	Q6WL78_STRAO	Q6wl78	streptomyce	512	20	95.2	113	1	RL34_METKA	RL34	methanopyru
440	20	95.2	101	2	Q6WL78_STRAO	Q6wl78	streptomyce	513	20	95.2	113	2	Q5OT98_ENTHI	Q5ot98	entamoeba h
441	20	95.2	101	2	Q6RR17_MYCOE	Q6rr17	mycobacteri	514	20	95.2	113	2	Q5YBC4_9CHLO	Q5ybc4	helicospori
442	20	95.2	101	2	Q6QMZ5_CHILA	Q6qmz5	chinchilla	515	20	95.2	113	2	Q5YBC4_9CHLO	Q5ybc4	helicospori
443	20	95.2	101	2	Q91G12_NPVEP	Q91g12	epiphyas po	516	20	95.2	113	2	Q94XC3_RHOMR	Q94xc3	rhodothermu
444	20	95.2	102	2	Q5XK30_CANAL	Q5xk30	candida alb	517	20	95.2	113	2	Q4UCV9_AGRV5	Q4ucv9	agrobacteri
445	20	95.2	102	2	Q78SG0_NEUCR	Q78sg0	neurospora	518	20	95.2	113	2	Q9HVK9_PSEAE	Q9hvk9	pseudomonas
446	20	95.2	102	2	Q701L6_9PSED	Q701l6	pseudomonas	519	20	95.2	114	2	Q50ZF8_ENTHI	Q50zf8	entamoeba h
447	20	95.2	102	2	Q88HT0_PSEPK	Q88ht0	pseudomonas	520	20	95.2	114	2	Q50ZF8_ENTHI	Q50zf8	entamoeba h
448	20	95.2	102	2	Q82O80_STRAW	Q82o80	streptomyce	521	20	95.2	114	2	Q7BF85_YERPS	Q7bf85	yersinia ps
449	20	95.2	103	2	Q9KLM2_VIBCH	Q9klm2	vibrio chol	522	20	95.2	114	2	Q7VMS3_BORPE	Q7vms3	bordetella
450	20	95.2	103	2	Q8BOK2_MOUSE	Q8bok2	mus musculu	523	20	95.2	114	2	Q7WJ54_BORBR	Q7wj54	bordetella
451	20	95.2	104	2	Q9U5Z1_CAERL	Q9u5z1	caenorhabdi	524	20	95.2	114	2	Q7WJ54_BORBR	Q7wj54	bordetella
452	20	95.2	104	2	Q5GXN0_XANOR	Q5gxno	xanthomonas	525	20	95.2	114	2	Q93PD5_YERPE	Q93pd5	yersinia pe
453	20	95.2	104	2	Q4K3L3_PSEFS	Q4k3l3	pseudomonas	526	20	95.2	114	2	Q66A07_YERPS	Q66a07	yersinia ps
454	20	95.2	104	2	Q8A641_BACTN	Q8a641	bacteroides	527	20	95.2	114	2	Q8NPD4_CORGL	Q8npd4	corynebacte
455	20	95.2	104	2	Q7ZZU0_OREMO	Q7zzu0	oreochromis	528	20	95.2	114	2	Q8KDU0_CHLTE	Q8kdu0	chlorobium
456	20	95.2	105	2	Q24309_PEA	Q24309	pisum sativ	529	20	95.2	114	2	Q8DG42_VIBVU	Q8dg42	vibrio vuln
457	20	95.2	105	2	Q88UH4_LACPL	Q88uh4	lactobacill	530	20	95.2	115	1	Q6T6J7_RAT	Q6t6j7	rattus norv
458	20	95.2	106	1	RL24_PORGI	Q7mtm4	porphyromon	531	20	95.2	115	2	YG32_YEAST	YG32	yeast
459	20	95.2	106	2	Q6CKM8_KULIA	Q6ckm8	kluyveromyc	532	20	95.2	115	2	Q518F0_ENTHI	Q518f0	entamoeba h
460	20	95.2	106	2	Q6NP30_DROME	Q6np30	drosophila	533	20	95.2	115	2	Q4V477_DROME	Q4v477	drosophila
461	20	95.2	106	2	Q9MYZ6_TRIVU	Q9myz6	trichosurus	534	20	95.2	115	2	Q4YH95_FLABM	Q4yh95	plasmodium
462	20	95.2	106	2	Q83EN0_COXBU	Q83en0	coxiella bu	535	20	95.2	115	2	Q4FUP3_9GAMM	Q4fup3	psychobact
463	20	95.2	107	2	Q58HJ1_SHEEP	Q58hj1	ovis aries	536	20	95.2	115	2	Q8ZIR6_YERPE	Q8zir6	yersinia pe
464	20	95.2	107	2	Q6WRV8_HAFAL	Q6wrv8	hafnia alve	537	20	95.2	115	2	Q66EX3_YERPS	Q66ex3	yersinia ps
465	20	95.2	107	2	Q6WRV5_HAFAL	Q6wr5	hafnia alve	538	20	95.2	115	2	Q62BN5_BURMA	Q62bn5	burkholderi
466	20	95.2	108	2	Q4FUK4_9GAMM	Q4fuk4	psychrobact	539	20	95.2	116	2	Q8NFC0_CORGL	Q8nfc0	corynebacte
467	20	95.2	108	2	Q4WI77_ASPFU	Q4wi77	aspergillus	540	20	95.2	116	2	Q8NFC0_CORGL	Q8nfc0	corynebacte
468	20	95.2	108	2	Q6Z2K8_ORYSA	Q6z2k8	oryza sativ	541	20	95.2	116	2	Q8NFC0_CORGL	Q8nfc0	corynebacte
469	20	95.2	108	2	Q9ZU61_ARYTH	Q9zu61	arabidopsis	542	20	95.2	116	2	Q8NFC0_CORGL	Q8nfc0	corynebacte
	20	95.2	108	2	Q81Y60_BACAN	Q81y60	bacillus an								

543	20	95.2	116	2	Q80JM6_9RHAB	Q80jm6 rabies_viru	616	20	95.2	125	2	Q936F9_STRAAU	Q936f9 staphylococ
544	20	95.2	116	2	Q80JM7_9RHAB	Q80jm7 rabies_viru	617	20	95.2	125	2	Q5GU44_XANOR	Q5gu44 xanthomonas
545	20	95.2	116	2	Q80JM8_9RHAB	Q80jm8 rabies_viru	618	20	95.2	125	2	Q5GB77_9HEPC	Q5gb77 hepatitis c
546	20	95.2	116	2	Q80JM9_9RHAB	Q80jm9 rabies_viru	619	20	95.2	125	2	Q9YLA3_9GEMI	Q9yla3 macropiliu
547	20	95.2	116	2	Q80JM0_9RHAB	Q80jm0 rabies_viru	620	20	95.2	126	1	URE2_BACPA	P41021 bacillus pa
548	20	95.2	116	2	Q80JN1_9RHAB	Q80jn1 rabies_viru	621	20	95.2	126	2	Q5V566_HALMA	Q5v566 haloarcula
549	20	95.2	116	2	Q80JN2_9RHAB	Q80jn2 rabies_viru	622	20	95.2	126	2	Q4J2G4_AZOVI	Q4j2g4 azotobacter
550	20	95.2	116	2	Q80JN3_9RHAB	Q80jn3 rabies_viru	623	20	95.2	126	2	Q6MPD7_BDEBA	Q6mpd7 bdellovibri
551	20	95.2	116	2	Q80JN4_9RHAB	Q80jn4 rabies_viru	624	20	95.2	127	2	Q8TX75_METKA	Q8tx75 methanoporu
552	20	95.2	116	2	Q80JN5_9RHAB	Q80jn5 rabies_viru	625	20	95.2	127	2	Q5E6E6_9CAUD	Q5e6e6 aeromonas p
553	20	95.2	116	2	Q80JN6_9RHAB	Q80jn6 rabies_viru	626	20	95.2	127	2	Q9L5Q7_SALTI	Q9l5q7 salmonella
554	20	95.2	116	2	Q41053_CHVP1	Q41053 paramedum	627	20	95.2	127	2	Q96622_9GEMI	Q96622 african tom
555	20	95.2	117	2	Q9Y6H7_HUMAN	Q9y6h7 homo sapien	628	20	95.2	128	2	Q9Y9Q1_AERPE	Q9y9q1 aeropyrum p
556	20	95.2	117	2	Q4N0X1_THEPA	Q4n0x1 theileria p	629	20	95.2	128	2	Q5AX09_EMENI	Q5ax09 aspergillus
557	20	95.2	117	2	Q7F0A4_ORYSA	Q7f0a4 oryza sativ	630	20	95.2	128	2	Q6S8S1_PLAFA	Q6s8s1 plasmodium
558	20	95.2	117	2	Q941S4_ORYSA	Q941s4 oryza sativ	631	20	95.2	128	2	Q4X8A8_PLACH	Q4x8a8 plasmodium
559	20	95.2	117	2	Q56GH4_PHYPA	Q56gh4 physcomitre	632	20	95.2	128	2	Q4XHD1_PLACH	Q4xhd1 plasmodium
560	20	95.2	117	2	Q8EL81_OCEIH	Q8el81 oceanobacil	633	20	95.2	128	2	Q04349_ARATH	Q04349 arabidopsis
561	20	95.2	118	1	ACPS_STRPY	Q99y97 streptococ	634	20	95.2	128	2	Q87154_VIBPA	Q87154 vibrio para
562	20	95.2	118	2	Q7CX77_AGR75	Q7cx77 agrobacteri	635	20	95.2	128	2	Q4TEC9_TETNG	Q4tec9 tetraodon n
563	20	95.2	118	2	Q5LNT4_SILPO	Q5lnt4 silicibacte	636	20	95.2	129	2	Q8U2C3_PYRFU	Q8u2c3 pyrococcus
564	20	95.2	118	2	Q88U14_LACPL	Q88u14 lactobacill	637	20	95.2	129	2	Q8U2P6_PYRFU	Q8u2p6 pyrococcus
565	20	95.2	118	2	Q4SBM5_TETNG	Q4sbm5 tetraodon n	638	20	95.2	129	2	Q58943_PYRHO	Q58943 pyrococcus
566	20	95.2	119	2	Q8LR81_ORYSA	Q8lr81 oryza sativ	639	20	95.2	129	2	Q4IF60_GIBZE	Q4if60 gibberella
567	20	95.2	119	2	Q5CCY7_9GEMI	Q5ccy7 tomato leaf	640	20	95.2	129	2	Q26382_LYMST	Q26382 lymanea sta
568	20	95.2	119	2	Q4W4Y3_9GEMI	Q4w4y3 tomato leaf	641	20	95.2	129	2	Q4ZMJ3_PSESY	Q4zmnj3 pseudomonas
569	20	95.2	120	2	Q9M714_9SOLA	Q9m714 petunia axi	642	20	95.2	129	2	Q5P129_AZOSE	Q5p129 azoarcus sp
570	20	95.2	120	2	Q8CL76_YERPE	Q8cl76 yerseinia pe	643	20	95.2	130	1	V29K_TRVTC	P05074 tobacco rat
571	20	95.2	120	2	Q5LE14_BACFN	Q5le14 bacteroides	644	20	95.2	130	2	Q9HKA4_THEAC	Q9hka4 thermoplasm
572	20	95.2	120	2	Q6P6E5_BRARE	Q6p6e5 brachydanio	645	20	95.2	130	2	Q6OL16_CAEER	Q6ol16 caenorhabdi
573	20	95.2	121	1	YXBE_BACSU	P54944 bacillus su	646	20	95.2	130	2	Q9N5E4_CAEEL	Q9n5e4 caenorhabdi
574	20	95.2	121	2	Q9YE40_AERPE	Q9ye40 aeropyrum p	647	20	95.2	130	2	Q912D5_PSEAE	Q912d5 pseudomonas
575	20	95.2	121	2	Q4ZVD1_PSESY	Q4zvd1 pseudomonas	648	20	95.2	130	2	Q62J42_BURMA	Q62j42 burkholderi
576	20	95.2	121	2	Q4ZYG7_PSESY	Q4zyg7 pseudomonas	649	20	95.2	130	2	Q87527_9RETR	P87527 bovine immu
577	20	95.2	121	2	Q6CYR3_ERWCT	Q6cyr3 erwinia car	650	20	95.2	130	2	P87529_9RETR	P87529 bovine immu
578	20	95.2	121	2	Q5CCM1_9GEMI	Q5ccm1 ageratum ye	651	20	95.2	130	2	P90289_9RETR	P90289 bovine immu
579	20	95.2	121	2	Q5CCW5_9GEMI	Q5ccw5 ageratum ye	652	20	95.2	131	2	Q64DL3_9ARCH	Q64dl3 uncultured
580	20	95.2	121	2	Q5CCW9_9GEMI	Q5ccw9 ageratum ye	653	20	95.2	131	2	Q62532_DROME	Q62532 drosophila
581	20	95.2	121	2	Q5CCX3_9GEMI	Q5ccx3 ageratum ye	654	20	95.2	131	2	Q04297_9DIPT	Q04297 scaptomyza
582	20	95.2	121	2	Q5CCX7_9GEMI	Q5ccx7 ageratum ye	655	20	95.2	131	2	Q5YA65_9CAUD	Q5ya65 bacillus cl
583	20	95.2	121	2	Q5CCZ5_9GEMI	Q5ccz5 ageratum ye	656	20	95.2	131	2	Q7Y0V7_ORYSA	Q7y0v7 oryza sativ
584	20	95.2	121	2	Q5DW23_9GEMI	Q5dw23 ageratum ye	657	20	95.2	131	2	Q6T1G0_RHIME	Q6t1g0 rhizobium m
585	20	95.2	122	1	M540_ARATH	P93309 arabidopsis	658	20	95.2	131	2	Q57TJ3_SALCH	Q57tj3 salmonella
586	20	95.2	122	1	RL35_CHICK	Q98tf7 gallus gall	659	20	95.2	131	2	Q5PIL6_SALPA	Q5pil6 salmonella
587	20	95.2	122	1	RL35_HIPCM	Q6uzf7 hippocampus	660	20	95.2	131	2	Q7CR90_SALTY	Q7cr90 salmonella
588	20	95.2	122	1	RL35_HUMAN	P42766 homo sapien	661	20	95.2	131	2	Q8XFZ2_SALTI	Q8xfz2 salmonella
589	20	95.2	122	1	RL35_OPHHA	Q69cj9 ophiophagu	662	20	95.2	131	2	Q8JVS0_9POTV	Q8jvs0 soybean mos
590	20	95.2	122	1	RL35_PIG	Q29361 sus scrofa	663	20	95.2	131	2	P90358_9RETR	P90358 bovine immu
591	20	95.2	122	1	RL35_PLAPE	Q5dvn6 platicthys	664	20	95.2	132	2	Q51RA4_MAGGR	Q51ra4 magnaporthe
592	20	95.2	122	1	RL35_XENTR	Q6pbc1 xenopus tro	665	20	95.2	132	2	Q6TM68_BPD31	Q6tm68 bacterioph
593	20	95.2	122	1	Q8SVW9_ENCCU	Q8svw9 encephalito	666	20	95.2	132	2	Q85712_BPMB2	Q85712 mycobacteri
594	20	95.2	122	2	Q6DR86_ARATH	Q6dr86 arabidopsis	667	20	95.2	132	2	Q6QET8_WYCCB	Q6qet8 mycobacteri
595	20	95.2	122	2	Q4FR01_9GAMM	Q4fr01 psychrobact	668	20	95.2	132	2	Q6QEU0_9MYCO	Q6qeu0 mycobacteri
596	20	95.2	122	2	Q7UM88_RHOBA	Q7um88 rhodopirell	669	20	95.2	132	2	Q5HWR5_CMJUR	Q5hwr5 campylobact
597	20	95.2	122	2	Q74CF4_GEOSL	Q74cf4 geobacter s	670	20	95.2	132	2	Q8EKD1_SHEON	Q8ekd1 shewanella
598	20	95.2	122	2	Q5CCY3_9GEMI	Q5ccy3 pepper yell	671	20	95.2	133	2	Q8EPV3_METMA	Q8epv3 methanosarc
599	20	95.2	122	2	Q5CCZ9_9GEMI	Q5ccz9 pepper yell	672	20	95.2	133	2	Q8TLA9_METAC	Q8tla9 methanosarc
600	20	95.2	122	2	Q5PPY9_XENLA	Q5ppy9 xenopus lae	673	20	95.2	133	2	Q4UCV2_THEAN	Q4ucv2 theileria a
601	20	95.2	123	2	Q4WKK4_ASPFU	Q4wkk4 aspergillus	674	20	95.2	133	2	Q9BK39_DRONE	Q9bk39 drosophila
602	20	95.2	123	2	Q8N8S8_HUMAN	Q8n8s8 homo sapien	675	20	95.2	133	2	Q81AM8_BACCR	Q81am8 bacillus ce
603	20	95.2	123	2	Q4VBY5_HUMAN	Q4vby5 homo sapien	676	20	95.2	133	2	Q88VV5_LACPL	Q88vv5 lactobacill
604	20	95.2	123	2	Q4ZSN0_PSESY	Q4zsn0 pseudomonas	677	20	95.2	134	2	Q45786_BACTU	Q45786 bacillus th
605	20	95.2	123	2	Q922F6_RHIME	Q922f6 rhizobium m	678	20	95.2	134	2	Q49007_WYCCA	Q49007 mycoplasma
606	20	95.2	123	2	Q5CCZ1_9GEMI	Q5ccz1 tomato leaf	679	20	95.2	134	2	Q8FAZ2_ECOL6	Q8faz2 escherichia
607	20	95.2	123	2	Q4TAF6_TETNG	Q4taf6 tetraodon n	680	20	95.2	134	2	Q92N40_RHIME	Q92n40 rhizobium m
608	20	95.2	123	2	Q4FZQ7_XENLA	Q4fzq7 xenopus lae	681	20	95.2	134	2	Q64G53_9HIV1	Q64g53 human immu
609	20	95.2	124	2	Q8WZJ6_SCHPO	Q8wzj6 schizosacch	682	20	95.2	135	2	Q86XT0_HUMAN	Q86xt0 homo sapien
610	20	95.2	124	2	Q8VLY5_THIFE	Q8vly5 thibacillu	683	20	95.2	135	2	Q9UDC2_HUMAN	Q9udc2 homo sapien
611	20	95.2	124	2	Q5LBE8_BACFN	Q5lbe8 bacteroides	684	20	95.2	135	2	Q5DFZ2_SCHJA	Q5dfz2 schistosoma
612	20	95.2	124	2	Q64RU6_BACFR	Q64ru6 bacteroides	685	20	95.2	135	2	Q17529_CAEEL	Q17529 caenorhabdi
613	20	95.2	124	2	Q4SGM6_TETNG	Q4sgm6 tetraodon n	686	20	95.2	135	2	Q949K4_LYCES	Q949k4 lycopersico
614	20	95.2	125	1	RCRB_AGR75	Q8ufc8 agrobacteri	687	20	95.2	135	2	P74277_SYNY3	P74277 synechocyst
615	20	95.2	125	2	Q8TQC6_METAC	Q8tqc6 methanosarc	688	20	95.2	135	2	Q8D3E9_WIGBR	Q8d3e9 wiggleswort

689	20	95.2	135	2	Q8R3H3_MOUSE	Q8r3h3 mus musculus	762	20	95.2	144	1	RS15_MESAU	P62842 mesocricetu
690	20	95.2	136	1	RL17_RICPR	Q9zct0 rickettsia	763	20	95.2	144	1	RS15_MOUSE	P62843 mus musculus
691	20	95.2	136	1	S25BP_HUMAN	Q9z266 homo sapien	764	20	95.2	144	1	RS15_PIG	P62844 sus scrofa
692	20	95.2	136	1	S25BP_MOUSE	Q9z266 mus musculus	765	20	95.2	144	1	RS15_RAT	P62845 rattus norv
693	20	95.2	136	1	S25BP_RAT	P60192 rattus norv	766	20	95.2	144	1	RS15_XENLA	P20342 xenopus lae
694	20	95.2	136	1	Q8PSY3_METMA	Q8psy3 methanosarc	767	20	95.2	144	1	RUVX_PSEAE	Q91699 pseudomonas
695	20	95.2	136	1	Q5SXU8_HUMAN	Q5sxu8 homo sapien	768	20	95.2	144	2	Q938K5_9CAUD	Q655q4 oryza sativ
696	20	95.2	136	2	Q4HH11_CAMCO	Q4hh11 campylobact	769	20	95.2	144	2	Q655O4_ORYSA	Q655q4 oryza sativ
697	20	95.2	136	2	Q62DY3_BURMA	Q62dy3 burkholderi	770	20	95.2	144	2	Q79XT6_STRP3	Q79xt6 streptococc
698	20	95.2	136	2	Q6CTN5_KUULA	Q6ctn5 kluyveromyce	771	20	95.2	144	2	Q5XAS5_STRP6	Q5xas5 streptococc
699	20	95.2	137	2	Q7YRK4_TARSY	Q7yrk4 tarsius syr	772	20	95.2	144	2	Q9A0N2_STRP8	Q9a0n2 streptococc
700	20	95.2	137	2	Q6G127_ECOLI	Q6g127 escherichia	773	20	95.2	144	2	Q8NZS2_STRP8	Q8nzs2 streptococc
701	20	95.2	137	2	Q8GAJ5_ARTNI	Q8gaj5 arthrobacte	774	20	95.2	144	2	Q7UT05_RHOBA	Q7ut05 rhodopirell
702	20	95.2	137	2	Q7ZM81_LEPIC	Q7zm81 leptospira	775	20	95.2	145	1	COX5A_MOUSE	P12787 mus musculus
703	20	95.2	137	2	Q8EYR4_LEPTOSPIRA	Q8eyr4 leptospira	776	20	95.2	145	1	RS15_XIPNA	P70066 xiphophorus
704	20	95.2	137	2	Q68WA3_RICKETTSIA	Q68wa3 rickettsia	777	20	95.2	145	1	RUVX_PSEFL	Q9f4i8 pseudomonas
705	20	95.2	138	2	Q5OZP0_ENTHI	Q5ozp0 entamoeba h	778	20	95.2	145	2	Q5BZ98_SCHJA	Q5bz98 schistosoma
706	20	95.2	138	2	Q88HL5_PSEPK	Q88hl5 pseudomonas	779	20	95.2	145	2	Q56K10_BOVIN	Q56k10 bos taurus
707	20	95.2	138	2	Q9TUU9_NEIMA	Q9tuh9 neisseria m	780	20	95.2	145	2	Q5RDI7_PONPY	Q5rdi7 pongo pygma
708	20	95.2	138	2	Q6T372_9PARA	Q6t372 newcastle d	781	20	95.2	145	2	Q4K4E4_PSEF5	Q4k4e4 pseudomonas
709	20	95.2	138	2	Q4RVV8_TETNG	Q4rvv8 tetraodon n	782	20	95.2	145	2	Q7UZN2_PROMP	Q7uzn2 prochloroco
710	20	95.2	139	1	RBS_OJILU	Q8tp87 methanosarc	783	20	95.2	145	2	Q6D612_ERWCT	Q6d6i2 erwinia car
711	20	95.2	139	2	Q8TP87_METAC	Q8tp87 methanosarc	784	20	95.2	145	2	Q7NPJ0_GLOVI	Q7npj0 gloeobacter
712	20	95.2	139	2	Q9TY82_SULSO	Q9ty82 sulfolobus	785	20	95.2	145	2	Q8FTB4_COREF	Q8ftb4 corynebacte
713	20	95.2	139	2	Q4YH82_PLABE	Q4yh82 plasmodium	786	20	95.2	145	2	Q63JB7_BURPS	Q63jb7 burkholderi
714	20	95.2	139	2	Q75NF9_9STRA	Q75nf9 heterosigma	787	20	95.2	145	2	Q52KC9_MOUSE	Q52kc9 mus musculus
715	20	95.2	139	2	Q6HVQ3_BACAN	Q6hqv3 bacillus an	788	20	95.2	145	2	Q8RJUC_XENTR	Q8rjuc xenopus tro
716	20	95.2	139	2	Q4MHW2_BACCE	Q4mhw2 bacillus ce	789	20	95.2	145	2	Q642P0_XENLA	Q642p0 xenopus lae
717	20	95.2	139	2	Q6HFR2_BACCH	Q6hfr2 bacillus th	790	20	95.2	145	2	Q6D6W6_BRARE	Q6dgw6 brachydanio
718	20	95.2	139	2	Q733V9_BACCI	Q733v9 bacillus ce	791	20	95.2	145	2	Q6Q421_BRARE	Q6q421 brachydanio
719	20	95.2	139	2	Q9KZZ9_STRCO	Q9kzz9 streptomyce	792	20	95.2	145	2	Q71U13_LAPHA	Q71u13 lapemis har
720	20	95.2	139	2	Q637T7_BACCC	Q637t7 bacillus ce	793	20	95.2	145	2	Q90YQ9 ICTPU	Q90yq9 ictalurus p
721	20	95.2	139	2	Q5L4H9_TYLCV	Q5l4h9 tomato yell	794	20	95.2	145	2	Q9YGQ0_SALSA	Q9ygq0 salmo salar
722	20	95.2	139	2	Q5L411_TYLCV	Q5l411 tomato yell	795	20	95.2	145	2	COX5A_RAT	P11240 rattus norv
723	20	95.2	139	2	Q5L413_TYLCV	Q5l413 tomato yell	796	20	95.2	146	1	Q8TJLI_METAC	Q8tjli methanosarc
724	20	95.2	139	2	Q634T6_9FLAV	Q634t6 tick-borne	797	20	95.2	146	2	Q5V2B3_HALMA	Q5v2b3 haloarcula
725	20	95.2	140	2	Q8TSR6_METAC	Q8tsr6 methanosarc	798	20	95.2	146	2	Q8ZSS3_PYRAE	Q8zss3 pyrobaculum
726	20	95.2	140	2	Q8ZYL8_PYRAE	Q8zyl8 pyrobaculum	799	20	95.2	146	2	Q9SHZ9_ARATH	Q9shz9 arabidopsis
727	20	95.2	140	2	Q8T7A5_AUTIO	Q8t7a5 automeris i	800	20	95.2	146	2	Q89YB2_BRAJH	Q89yb2 bradyrhizob
728	20	95.2	140	2	Q9MCK1_9CAUD	Q9mck1 streptococc	801	20	95.2	146	2	Q33703_STRPN	Q33703 streptococc
729	20	95.2	140	2	Q9XJA4_9CAUD	Q9xja4 streptococc	802	20	95.2	146	2	Q9D2W1_MOUSE	Q9d2w1 mus musculus
730	20	95.2	140	2	Q8S3W2_CUCSA	Q8s3w2 cucumis sat	803	20	95.2	147	2	Q9RW8_NEUCR	Q9rw8 neuropept
731	20	95.2	140	2	Q5LED0_BACFN	Q5led0 bacteroides	804	20	95.2	147	2	Q5CEJ1_CRYHO	Q5cej1 cryptospori
732	20	95.2	140	2	Q64VI3_BACFR	Q64vi3 bacteroides	805	20	95.2	147	2	Q61585_ORYSA	Q61585 oryza sativ
733	20	95.2	140	2	Q88ML7_PSEPK	Q88ml7 pseudomonas	806	20	95.2	147	2	Q8DNF8_STRG6	Q8dnf8 streptococc
734	20	95.2	140	2	Q5LW31_SILPO	Q5lw31 silicibacte	807	20	95.2	147	2	Q97NS3_STRPN	Q97ns3 streptococc
735	20	95.2	140	2	Q8X8S0_ECO57	Q8x8s0 escherichia	808	20	95.2	147	2	Q6N6L2_RHOPA	Q6n6l2 rhodopaeu
736	20	95.2	140	2	Q83LA5_SHIFL	Q83la5 shigella fl	809	20	95.2	147	2	Q7URA3_RHOBA	Q7ura3 rhodopirell
737	20	95.2	140	2	Q83LQ6_SHIFL	Q83lq6 shigella fl	810	20	95.2	147	1	YHT6_YEAST	P38839 saccharomyc
738	20	95.2	141	1	RUVX_PSEPK	Q88d32 pseudomonas	811	20	95.2	148	2	P87290_YEAST	P87290 saccharomyc
739	20	95.2	141	2	Q8T7A0_9NEOP	Q8t7a0 eaeles impe	812	20	95.2	148	2	Q706P2_PSEPU	Q706p2 pseudomonas
740	20	95.2	141	2	Q9BNZ3_DRYVU	Q9bnz3 dryocampa r	813	20	95.2	148	2	Q5H6P1_XANOR	Q5h6p1 xanthomonas
741	20	95.2	141	2	Q64278_9CAUD	Q64278 streptococc	814	20	95.2	148	2	Q4UV34_XANCP	Q4uv34 xanthomonas
742	20	95.2	141	2	Q4K9C8_PSEF5	Q4k9c8 pseudomonas	815	20	95.2	148	2	Q6IP26_XENLA	Q6ip26 xenopus lae
743	20	95.2	141	2	Q6LRD8_PHOPR	Q6lrd8 photobacter	816	20	95.2	148	2	P87528_9RETR	P87528 bovine immu
744	20	95.2	141	2	Q71S22_9SCOM	Q71s22 scombridae	817	20	95.2	148	2	P87530_9RETR	P87530 bovine immu
745	20	95.2	141	2	Q71S23_CYPCA	Q71s23 cyprinus ca	818	20	95.2	149	1	Y1039_EORPE	Q7vr88 bordetella
746	20	95.2	141	2	Q4RS92_TETNG	Q4rs92 tetraodon n	819	20	95.2	149	2	Q5CIG8_SCHJA	Q5cig8 schistosoma
747	20	95.2	142	2	Q50TF4_ENTHI	Q50tf4 entamoeba h	820	20	95.2	149	2	Q5QC93_AEDAE	Q5qc93 aedes aegyp
748	20	95.2	142	2	Q5YEW2_FRAAN	Q5yew2 fragaria an	821	20	95.2	149	2	Q5MIR4_AEDAL	Q5mir4 aedes albop
749	20	95.2	142	2	Q82FS2_STRAW	Q82fs2 streptomyce	822	20	95.2	149	2	Q6S7F8_9BACT	Q6s7f8 uncultured
750	20	95.2	142	2	Q90153_9FLAV	Q90153 hepaticis g	823	20	95.2	149	2	Q6S7P1_9BACT	Q6s7p1 uncultured
751	20	95.2	143	2	Q6MYB1_ASPPF	Q6myb1 aspergillus	824	20	95.2	149	2	Q6S7P5_9BACT	Q6s7p5 uncultured
752	20	95.2	143	2	Q584P6_9TRYP	Q584p6 trypanosoma	825	20	95.2	149	2	Q6S7P8_9BACT	Q6s7p8 uncultured
753	20	95.2	143	2	Q851L7_ORYSA	Q851l7 oryza sativ	826	20	95.2	149	2	Q6S7P9_9BACT	Q6s7p9 uncultured
754	20	95.2	143	2	Q9FW54_ORYSA	Q9fw54 oryza sativ	827	20	95.2	149	2	Q6S7Q2_9BACT	Q6s7q2 uncultured
755	20	95.2	143	2	Q53UK5_ORYSA	Q53jk5 oryza sativ	828	20	95.2	149	2	Q6S7Q9_9BACT	Q6s7q9 uncultured
756	20	95.2	143	2	Q8GBU6_9BACT	Q8gbu6 uncultured	829	20	95.2	149	2	Q8RAH8_THETN	Q8rah8 thermonaer
757	20	95.2	143	2	Q5H445_XANOR	Q5h445 xanthomonas	830	20	95.2	149	2	Q9TDW3_CLOAB	Q9tdw3 clostridium
758	20	95.2	143	2	Q9AK79_STRCO	Q9ak79 streptomyce	831	20	95.2	149	2	Q4S8B4_TETNG	Q4s8b4 tetraodon n
759	20	95.2	144	1	RPI_ECOLI	P13809 escherichia	832	20	95.2	150	1	COX5A_HUMAN	P20674 homo sapien
760	20	95.2	144	1	RS15_CHICK	P62846 gallus gall	833	20	95.2	150	2	Q8TB65_MELLI	Q8tb65 melampora
761	20	95.2	144	1	RS15_HUMAN	P62841 homo sapien	834	20	95.2	150	2	Q8TB65_HUMAN	Q8tb65 homo sapien



835	20	95.2	150	2	Q5DBT0_SCHJA	Q5dbt0 schistoeoma	908	20	95.2	156	2	Q4J347_AZOV1	Q4j347 azobacter
836	20	95.2	150	2	Q4Z529_PLABE	Q4z529 plasmodium	909	20	95.2	156	2	Q87A89_XYLFT	Q87a89 xylella fas
837	20	95.2	150	2	Q53CF8_MACMU	Q53cf8 macaca mula	910	20	95.2	156	2	Q8A556_BACTN	Q8a556 bacteroides
838	20	95.2	150	2	Q53CG1_SAIISC	Q53cg1 saimiri sci	911	20	95.2	157	2	Q88A24_PSESM	Q88a24 pseudomonas
839	20	95.2	150	2	Q5HT49_CAMJR	Q5ht49 campylobact	912	20	95.2	157	1	Y1139_BORPA	Q7wb68 bordetella
840	20	95.2	150	2	Q9PWT8_CAMJE	Q9pwt8 campylobact	913	20	95.2	157	1	Y1355_BORBR	Q7wn5 bordetella
841	20	95.2	150	2	Q884K7_PSESM	Q884k7 pseudomonas	914	20	95.2	157	2	Q9VQL3_DROME	Q9vql3 drosophila
842	20	95.2	150	2	Q67NK6_SYMTH	Q67nk6 symbiobacte	915	20	95.2	157	2	Q6ZDG7_ORISA	Q6zdg7 oryza sativ
843	20	95.2	150	2	Q8Q976_9CALI	Q8q976 reptile cal	916	20	95.2	157	2	Q6ZKY5_ORISA	Q6zky5 oryza sativ
844	20	95.2	150	2	Q094N0_SIVCZ	Q094n0 chimpanzee	917	20	95.2	157	2	Q5F7S5_NEIG1	Q5f7s5 neisseria g
845	20	95.2	151	1	RUVX_NEIG1	Q5f936 neisseria g	918	20	95.2	157	2	Q7NPf8_GLOVI	Q7npf8 gloebacter
846	20	95.2	151	1	RUVX_NEIMA	Q9j11 neisseria m	919	20	95.2	157	2	Q7V7E2_PROWM	Q7v7e2 prochloroco
847	20	95.2	151	1	RUVX_NEIMB	Q9j216 neisseria m	920	20	95.2	158	2	Q55CP5_DICDI	Q55cp5 dictyosteli
848	20	95.2	151	1	Y903_AQUAE	Q67050 aquifex aeo	921	20	95.2	158	2	Q5TWC0_ANOGA	Q5twc0 anopheles g
849	20	95.2	151	2	Q4WS67_ASFPU	Q4ws67 aspergillus	922	20	95.2	158	2	Q6YDF3_9FABA	Q6ydf3 arachis car
850	20	95.2	151	2	Q5S1X4_IXOSC	Q5s1x4 ixodes scap	923	20	95.2	158	2	Q6YWI6_ORISA	Q6ywi6 oryza sativ
851	20	95.2	151	2	Q25015_9VEST	Q25015 haliotis au	924	20	95.2	158	2	Q5GUC0_XANOR	Q5guc0 xanthomonas
852	20	95.2	151	2	Q9KKH6_YEREN	Q9kkh6 yersinia en	925	20	95.2	158	2	Q5GVP4_XANOR	Q5gvp4 xanthomonas
853	20	95.2	151	2	Q6XA36_9PARA	Q6xa36 newcastle d	926	20	95.2	158	2	Q5H030_XANOR	Q5h030 xanthomonas
854	20	95.2	151	2	Q6XA38_9PARA	Q6xa38 newcastle d	927	20	95.2	158	2	Q5H119_XANOR	Q5h119 xanthomonas
855	20	95.2	151	2	Q6XA40_9PARA	Q6xa40 newcastle d	928	20	95.2	158	2	Q5H3K2_XANOR	Q5h3k2 xanthomonas
856	20	95.2	151	2	Q6XA42_9PARA	Q6xa42 newcastle d	929	20	95.2	158	2	Q5H3S0_XANOR	Q5h3s0 xanthomonas
857	20	95.2	151	2	Q6XA44_9PARA	Q6xa44 newcastle d	930	20	95.2	158	2	Q5H648_XANOR	Q5h648 xanthomonas
858	20	95.2	151	2	Q6XA46_9PARA	Q6xa46 newcastle d	931	20	95.2	158	2	Q6MPD6_BDRBA	Q6mpd6 bdellovibri
859	20	95.2	151	2	Q6XA48_9PARA	Q6xa48 newcastle d	932	20	95.2	159	2	Q6BZQ8_YARLI	Q6bzq8 yarrowia li
860	20	95.2	151	2	Q6XA50_9PARA	Q6xa50 newcastle d	933	20	95.2	159	2	Q5BXJ9_SCHJA	Q5bxj9 schistosoma
861	20	95.2	151	2	Q6XA52_9PARA	Q6xa52 newcastle d	934	20	95.2	159	2	Q4YH56_PLABE	Q4yh56 plasmodium
862	20	95.2	151	2	Q6XA54_9PARA	Q6xa54 newcastle d	935	20	95.2	159	2	Q6VTB3_9BACT	Q6vtb3 symbiont ba
863	20	95.2	151	2	Q6XA56_9PARA	Q6xa56 newcastle d	936	20	95.2	159	2	Q4LBL9_PSESH	Q4lbl9 pseudomonas
864	20	95.2	151	2	Q6XA58_9PARA	Q6xa58 newcastle d	937	20	95.2	159	2	Q4KD78_PSEF5	Q4kd78 pseudomonas
865	20	95.2	151	2	Q6XA60_9PARA	Q6xa60 newcastle d	938	20	95.2	159	2	Q5M419_SYNFP	Q5m419 synchococc
866	20	95.2	151	2	Q6XA62_9PARA	Q6xa62 newcastle d	939	20	95.2	159	2	Q5Z3B1_NOCPA	Q5z3b1 nocardia fa
867	20	95.2	151	2	Q6NYP1_BRARE	Q6nyp1 brachydanio	940	20	95.2	159	2	Q6PHI9_XANAC	Q6phi9 xanthomonas
868	20	95.2	151	2	Q6PEX0_BRARE	Q6pex0 brachydanio	941	20	95.2	159	2	Q6XA26_9PARA	Q6xa26 newcastle d
869	20	95.2	151	2	Q4T9H6_TETNG	Q4t9h6 tetraodon n	942	20	95.2	159	2	Q6XA28_9PARA	Q6xa28 newcastle d
870	20	95.2	151	2	Q4SMA7_TETNG	Q4sma7 tetraodon n	943	20	95.2	159	2	Q6XA30_9PARA	Q6xa30 newcastle d
871	20	95.2	152	1	COX5A_BOVIN	P00426 bos taurus	944	20	95.2	159	2	Q6XA34_9PARA	Q6xa34 newcastle d
872	20	95.2	152	2	Q4IH16_GIBZE	Q4ih16 gibberella	945	20	95.2	160	1	PHAA_AGLNE	P28555 aglaotahamni
873	20	95.2	152	2	Q7PSR9_ANOGA	Q7psr9 anopheles g	946	20	95.2	160	1	Y650_TREPA	Q83656 treponema p
874	20	95.2	152	2	Q61J02_DROME	Q61j02 drosophila	947	20	95.2	160	2	Q8PXK2_METMA	Q8pxk2 methanosaer
875	20	95.2	152	2	Q50M98_ENTHI	Q50m98 entamoeba h	948	20	95.2	160	2	Q75EX7_ASHGO	Q75ex7 ashbya goss
876	20	95.2	152	2	Q5XD03_STRP6	Q5xd03 streptococc	949	20	95.2	160	2	Q7PN92_ANOGA	Q7pn92 anopheles g
877	20	95.2	152	2	Q9A0K8_STRPY	Q9a0k8 streptococc	950	20	95.2	160	2	Q8WZS9_ORISA	Q8wz89 oryza sativ
878	20	95.2	152	2	Q745S5_MYCPA	Q745s5 mycobacteri	951	20	95.2	160	2	Q8H2S9_ORISA	Q8h2s9 xanthomonas
879	20	95.2	152	2	Q7TVS1_MYCBO	Q7tvs1 mycobacteri	952	20	95.2	160	2	Q5H107_XANOR	Q5h107 xanthomonas
880	20	95.2	152	2	Q8P1M2_STRP8	Q8p1m2 streptococc	953	20	95.2	160	2	Q6F6S8_ACIAD	Q6f6s8 acinetobact
881	20	95.2	152	2	Q69719_MYCTU	Q69719 mycobacteri	954	20	95.2	160	2	Q7U4N4_SYNPX	Q7u4n4 synchococc
882	20	95.2	152	2	Q7CF99_STRP3	Q7cf99 streptococc	955	20	95.2	161	1	MLR_DICDI	Q6tnq4 brachydanio
883	20	95.2	153	2	Q6B867_9ACAR	Q6b867 ixodes paci	956	20	95.2	161	2	Q5S2B5_DICDI	P13933 dictyosteli
884	20	95.2	153	2	Q9Y1I7_9VEST	Q9y1i7 tegula aure	957	20	95.2	161	2	Q5Q9W9_NICGL	Q5q9w9 nicotiana g
885	20	95.2	153	2	Q61LV3_DROME	Q61lv3 drosophila	958	20	95.2	161	2	Q4KB46_PSEF5	Q4kb46 pseudomonas
886	20	95.2	153	2	Q50VS3_ENTHI	Q50vs3 entamoeba h	959	20	95.2	161	2	Q9ZKH9_HELPJ	Q9zkh9 helicobacte
887	20	95.2	153	2	Q64289_9CAUD	Q64289 streptococc	960	20	95.2	162	1	CAVZ_CANPA	Q46550 canis famil
888	20	95.2	153	2	Q5FTK5_GLOUX	Q5ftk5 gluconobact	961	20	95.2	162	2	Q51A31_ENTHI	Q51a31 entamoeba h
889	20	95.2	153	2	Q7UB16_SYNPX	Q7ub16 synchococc	962	20	95.2	162	2	Q4Q087_LEIMA	Q4q087 leishmania
890	20	95.2	154	1	Y726_AZOSE	Q5p762 azoarcus sp	963	20	95.2	162	2	Q4FD05_PIG	Q4fd05 sus scrofa
891	20	95.2	154	2	Q61VP3_CABBR	Q61vp3 caenorhabdi	964	20	95.2	162	2	Q69752_PSEAE	Q69752 pseudomonas
892	20	95.2	154	2	Q5C1K9_SCHJA	Q5c1k9 schistosoma	965	20	95.2	162	2	Q9HWH1_PSEAE	Q9hwh1 pseudomonas
893	20	95.2	154	2	Q8GMN3_AERSA	Q8gm3 aeromonas s	966	20	95.2	162	2	Q5NYT4_AZOSE	Q5nyt4 azoarcus sp
894	20	95.2	154	2	Q5P0G9_AZOSE	Q5p0g9 azoarcus sp	967	20	95.2	162	2	Q6Z5H2_PERLE	Q6z5h2 peromyescus
895	20	95.2	154	2	Q8EMU1_OCEIH	Q8emu1 oceanobacil	968	20	95.2	162	2	Q6ZMQ5_MOUSE	Q6zmq5 mus musculu
896	20	95.2	154	2	Q6T343_9PARA	Q6t343 newcastle d	969	20	95.2	162	2	Q5ZHT0_CHICK	Q5zht0 gallus gall
897	20	95.2	154	2	Q6T370_9PARA	Q6t370 newcastle d	970	20	95.2	162	2	Q7ZV55_XENLA	Q7zv55 xenopus lae
898	20	95.2	155	2	Q4XEJ2_PLACH	Q4xej2 plasmodium	971	20	95.2	162	2	Q803J1_BRARE	Q803j1 brachydanio
899	20	95.2	155	2	Q7VS68_BORPE	Q7vs68 bordetella	972	20	95.2	162	2	Q5EBB4_XENTR	Q5eb4 xenopus tro
900	20	95.2	155	2	Q8XPQ0_RALSO	Q8xpq0 ralatonia s	973	20	95.2	162	2	Q6GPZ0_XENLA	Q6gpz0 xenopus lae
901	20	95.2	156	1	ELYS_HALCR	Q01380 haliotis cr	974	20	95.2	162	2	Q4RAF8_TETNG	Q4raf8 tetraodon n
902	20	95.2	156	2	Q55W00_CRYNE	Q55w00 cryptococcu	975	20	95.2	163	2	Q5DHH7_SCHJA	Q5dhh7 schistosoma
903	20	95.2	156	2	Q5KK91_CRYNE	Q5kk91 cryptococcu	976	20	95.2	163	2	Q9W372_DROME	Q9w372 drosophila
904	20	95.2	156	2	Q7PFI1_PLAYO	Q7pfi1 plasmodium	977	20	95.2	163	2	Q4XQC8_PLACH	Q4xc8 plasmodium
905	20	95.2	156	2	Q7P329_FUSNV	Q7p329 fusobacteri	978	20	95.2	163	2	Q9XUF0_CABEL	Q9xuf0 caenorhabdi
906	20	95.2	156	2	Q5GV41_XANOR	Q5gv41 xanthomonas	979	20	95.2	163	2	Q9FPV3_PETIN	Q9fpv3 petunia int
907	20	95.2	156	2	Q5GZ78_XANOR	Q5gz78 xanthomonas	980	20	95.2	163	2	Q8DG89_SYNEL	Q8dg89 synchococc

981 20 95.2 164 2 Q8N112 HUMAN  
982 20 95.2 164 2 Q6EPZ1\_ORYSA  
983 20 95.2 164 2 Q7XHL4\_ORYSA  
984 20 95.2 164 2 Q8XSK9\_RALSO  
985 20 95.2 164 2 Q9KKX4\_VIBCH  
986 20 95.2 164 2 Q67RM1\_SYWTH  
987 20 95.2 164 2 Q5F228\_MOUSE  
988 20 95.2 164 2 Q5RH22\_BRARE  
989 20 95.2 165 1 SNX12\_MOUSE  
990 20 95.2 165 2 Q513T7\_ENTHI  
991 20 95.2 165 2 Q4YIV0\_PLABE  
992 20 95.2 165 2 Q71IT5\_LACDL  
993 20 95.2 165 2 Q4J2V2\_AZOV1  
994 20 95.2 165 2 Q8D7T9\_VIBVU  
995 20 95.2 165 2 Q928B3\_LISIN  
996 20 95.2 165 2 Q986B3\_RHILO  
997 20 95.2 166 2 Q9BK38\_9DIPT  
998 20 95.2 166 2 Q56C19\_9CAUD  
999 20 95.2 166 2 Q8GQ00\_PSEAE  
1000 20 95.2 166 2 Q8DN14\_STRRG

## ALIGNMENTS

RESULT 1  
Q23912 DICDI PRELIMINARY; PRT; 16 AA.  
AC Q23912;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Nucleoside diphosphate kinase Gipl7 (Fragment).  
OS Dictyostelium discoideum (Slime mold).  
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.  
OX NCBI\_TaxID=44689;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=90277608; PubMed=2161830;  
RA Lacombe M.-L., Wallat V., Troll H., Veron M.;  
RT "Functional cloning of a nucleoside diphosphate kinase from  
RT Dictyostelium discoideum";  
RL J. Biol. Chem. 265:10012-10018(1990).  
DR EMBL; M36679; AAA33232.1; -; mRNA.  
DR Dictybase; DDB0185051; ndKC.  
DR GO; GO:0016301; F:Kinase activity; IEA.  
KW Kinase.  
FT NON\_TER 16 16  
SQ SEQUENCE 16 AA; 1897 MW; 38BE6B475C59885E CRC64;

Query Match 95.2%; Score 20; DB 2; Length 16;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
DB 6 RRLN 9

RESULT 2  
Q5BX77\_SCHJA PRELIMINARY; PRT; 27 AA.  
AC Q5BX77;  
DT 10-MAY-2005 (TrEMBLrel. 30, Created)  
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)  
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)  
DE Hypothetical protein.  
OS Schistosoma japonicum (Blood fluke).  
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;  
OC Schistosomatoida; Schistosomatidae; Schistosoma.  
OX NCBI\_TaxID=6182;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.

RA Han Z.;  
RL Submitted (MAR-2005) to the EMBL/GenBank/DBSJ databases.  
DR EMBL; AY812109; AAX27998.1; -; mRNA.  
DR InterPro; IPR011989; ARM-like.  
KW Hypothetical protein.  
SQ SEQUENCE 27 AA; 3410 MW; 7E0683D08C942315 CRC64;  
Query Match 95.2%; Score 20; DB 2; Length 27;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RRLN 4  
DB 6 RRLN 9

## RESULT 3

Q50L67\_DROBP PRELIMINARY; PRT; 33 AA.  
AC Q50L67;  
DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
DE Hypothetical protein CG11779 (Fragment).  
GN Name=CG11779;  
OS Drosophila bipectinata (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=42026;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=Wild type B133, and wild type CJB162;  
RA Nozawa M., Aotsuka T., Tamura K.;  
RT "A novel chimeric gene, sirene, found in the Drosophila bipectinata  
RT species complex: potential of retroposition with regulatory  
RT sequence";  
RL Submitted (NOV-2004) to the EMBL/GenBank/DBSJ databases.  
DR EMBL; AB194439; BAD98229.1; -; Genomic\_DNA.  
DR EMBL; AB194442; BAD98232.1; -; Genomic\_DNA.  
DR EMBL; AB194429; BAD98204.1; -; Genomic\_DNA.  
KW Hypothetical protein.  
FT NON\_TER 33 33  
SQ SEQUENCE 33 AA; 4154 MW; B9FF28FCB8AB796C CRC64;  
Query Match 95.2%; Score 20; DB 2; Length 33;  
Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
DB 25 RRLN 28

## RESULT 4

Q50L79\_9DIPT PRELIMINARY; PRT; 33 AA.  
AC Q50L79;  
DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
DE Hypothetical protein CG11779 (Fragment).  
GN Name=CG11779;  
OS Drosophila pseudoananassae pseudoananassae.  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=65965;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Nozawa M., Aotsuka T., Tamura K.;  
RT "A novel chimeric gene, sirene, found in the Drosophila bipectinata  
RT species complex: potential of retroposition with regulatory

RT sequence.";  
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB194433; BAD98216.1; -; Genomic\_DNA.  
KW Hypothetical protein.  
FT NON\_TER 33 33  
SQ SEQUENCE 33 AA; 4128 MW; E9FF28FCB8B686D CRC64;

Query Match 95.2%; Score 20; DB 2; Length 33;  
Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
|||  
Db 25 RRLN 28

## RESULT 5

Q50L83\_9DIPT  
ID Q50L83\_9DIPT PRELIMINARY; PRT; 33 AA.  
AC Q50L83;  
DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
DE Hypothetical protein CG11779 (Fragment).  
GN Name=CG11779;  
OS Drosophila pseudoscutellaris nigrens.  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=296643;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Nozawa M., Aotsuka T., Tamura K.;  
RT "A novel chimeric gene, sirene, found in the Drosophila bipectinata  
RT species complex: potential of retroposition with regulatory  
RT sequence.";  
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB194432; BAD98213.1; -; Genomic\_DNA.  
KW Hypothetical protein.  
FT NON\_TER 33 33  
SQ SEQUENCE 33 AA; 4128 MW; E9FF28FCB8B686D CRC64;

Query Match 95.2%; Score 20; DB 2; Length 33;  
Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
|||  
Db 25 RRLN 28

## RESULT 6

Q50L86\_9DIPT  
ID Q50L86\_9DIPT PRELIMINARY; PRT; 33 AA.  
AC Q50L86;  
DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
DE Hypothetical protein CG11779 (Fragment).  
GN Name=CG11779;  
OS Drosophila malerkotliana pallens.  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=296644;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Nozawa M., Aotsuka T., Tamura K.;  
RT "A novel chimeric gene, sirene, found in the Drosophila bipectinata  
RT species complex: potential of retroposition with regulatory  
RT sequence.";  
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB194431; BAD98210.1; -; Genomic\_DNA.

KW Hypothetical protein.  
FT NON\_TER 33 33  
SQ SEQUENCE 33 AA; 4154 MW; E9FF28FCB8AB796C CRC64;

Query Match 95.2%; Score 20; DB 2; Length 33;  
Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
|||  
Db 25 RRLN 28

## RESULT 7

Q50L89\_9DIPT  
ID Q50L89\_9DIPT PRELIMINARY; PRT; 33 AA.  
AC Q50L89;  
DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
DE Hypothetical protein CG11779 (Fragment).  
GN Name=CG11779;  
OS Drosophila malerkotliana malerkotliana.  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=195057;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Nozawa M., Aotsuka T., Tamura K.;  
RT "A novel chimeric gene, sirene, found in the Drosophila bipectinata  
RT species complex: potential of retroposition with regulatory  
RT sequence.";  
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB194430; BAD98207.1; -; Genomic\_DNA.  
KW Hypothetical protein.  
FT NON\_TER 33 33  
SQ SEQUENCE 33 AA; 4154 MW; E9FF28FCB8AB796C CRC64;

Query Match 95.2%; Score 20; DB 2; Length 33;  
Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
|||  
Db 25 RRLN 28

## RESULT 8

Q50L77\_DROAN  
ID Q50L77\_DROAN PRELIMINARY; PRT; 34 AA.  
AC Q50L77;  
DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
DE Hypothetical protein CG11779 (Fragment).  
GN Name=CG11779;  
OS Drosophila ananassae (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7217;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Nozawa M., Aotsuka T., Tamura K.;  
RT "A novel chimeric gene, sirene, found in the Drosophila bipectinata  
RT species complex: potential of retroposition with regulatory  
RT sequence.";  
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB194434; BAD98219.1; -; Genomic\_DNA.  
KW Hypothetical protein.  
FT NON\_TER 34 34  
SQ SEQUENCE 34 AA; 4262 MW; 90CF599A9912FA05 CRC64;

Query Match 95.2%; Score 20; DB 2; Length 34;  
 Best Local Similarity 100.0%; Pred. No. 4.1e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
 DB 26 RRLN 29

## RESULT 9

Q4ZG64 HUMAN PRELIMINARY; PRT; 35 AA.  
 AC Q4ZG64;  
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
 DE Hypothetical protein SP100 (Fragment).  
 GN Name=SP100;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
 OC Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Abbott A., McLellan M., Haub K.;  
 RT "The sequence of Homo sapiens BAC clone RP11-69J7.";  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Waterston R.H.;  
 RL Waterston R.H.;  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Waterston R.;  
 RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Wilson R.K.;  
 RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC009949; AX88870.1; -; Genomic\_DNA.  
 KW Hypothetical protein.  
 FT NON\_TER 35  
 SQ SEQUENCE 35 AA; 3729 MW; 0DD1F466F6016B22 CRC64;

Query Match 95.2%; Score 20; DB 2; Length 35;  
 Best Local Similarity 100.0%; Pred. No. 4.2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
 DB 11 RRLN 14

## RESULT 10

Q888U8 PSESM PRELIMINARY; PRT; 35 AA.  
 AC Q888U8;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Hypothetical protein.  
 GN OrderedLocusNames=PSPT00918;  
 OS Pseudomonas syringae (pv. tomato).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 OX NCBI\_TaxID=323;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=DC3000;  
 RX MEDLINE=22834015; PubMed=12928499; DOI=10.1073/pnas.1731982100;  
 RA Buell C.R., Joardar V., Lindeberg M., Selengut J., Paulsen I.T.,

RA Gwinn M.L., Dodson R.J., DeBoy R.T., Durkin A.S., Kolonay J.F.,  
 RA Madupu R., Daugherty S.C., Brinkac L.M., Beanan M.J., Haft D.H.,  
 RA Nelson W.C., Davidson T.M., Zafar N., Zhou L., Liu J., Yuan Q.,  
 RA Khouri H.M., Fedorova N.B., Tran B., Russell D., Herry K.J.,  
 RA Utterback T.R., Van Aken S.E., Feldblyum T.V., D'Ascenzo M.,  
 RA Deng W.-L., Ramos A.R., Alfano J.R., Cartinhour S., Chatterjee A.K.,  
 RA Delaney T.P., Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X.,  
 RA Bender C.L., White O., Fraser C.M., Collmer A.;  
 RT "The complete genome sequence of the Arabidopsis and tomato pathogen  
 RT Pseudomonas syringae pv. tomato DC3000.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186(2003).  
 DR EMBL; AE016853; AA054452.1; -; Genomic\_DNA.  
 DR TIGR; PSPT00918; -;  
 KW Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 35 AA; 4182 MW; 6F44D83F7F895634 CRC64;

Query Match 95.2%; Score 20; DB 2; Length 35;  
 Best Local Similarity 100.0%; Pred. No. 4.2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
 DB 3 RRLN 6

## RESULT 11

Q4SVAL TETNG PRELIMINARY; PRT; 35 AA.  
 AC Q4SVAL;  
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
 DE Chromosome undetermined SCAP13770, whole genome shotgun sequence.  
 DE (Fragment).  
 GN ORFNames=GSTENG00012082001;  
 OS Tetraodon nigroviridis (Green puffer).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
 OC Tetraodontidae; Tetraodontidae; Tetraodon.  
 OX NCBI\_TaxID=99883;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,  
 RA Mauceli E., Bouneau L., Fischer C., Ozou-Costaz C., Bernot A.,  
 RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,  
 RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,  
 RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,  
 RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,  
 RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J., Bosak S.,  
 RA Parra G., Lardier G., Chappie C., McKernan K.J., McEwan P., Mesirov J.,  
 RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,  
 RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,  
 RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,  
 RA Wincker P., Lander E.S., Weissbach J., Roest Crolius H.;  
 RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals  
 RT the early vertebrate proto-karyotype.";  
 RL Nature 431:946-957(2004).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Genoscope; Whitehead Institute Centre for Genome Research;  
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
 CC -1- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; CAAE01013770; CAF95431.1; -; Genomic\_DNA.  
 FT NON\_TER 1  
 SQ SEQUENCE 35 AA; 3917 MW; 104DEC79CCD12A70 CRC64;

Query Match 95.2%; Score 20; DB 2; Length 35;  
 Best Local Similarity 100.0%; Pred. No. 4.2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
DB 26 RRLN 29

Best Local Similarity 100.0%; Pred. No. 4.3e+02; Mismatches 0; Indels 0; Gaps 0;

Query Match 95.2%; Score 20; DB 2; Length 36;

Q425P5 PLABE PRELIMINARY; PRT; 36 AA.

DT 13-SEP-2005 (TrEMBLrel. 31, Created)

DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)

DE Hypothetical protein (Fragment).

GN ORENames=PB100986.00.0;

OS Plasmodium berghei.

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI\_TaxID=5821;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RA Hall N., Karras M., Raine J.D., Carlton J.M., Kooij T.W.A., Berriman M., Florens D., Janssen C.S., Pain A., Christophides G.K., James K., Rutherford K., Harris B., Harris H.E., Churcher C., Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J., Bidwell S.L., Rajandream M.A., Carucci D.J., Yates J.R., Kafatos F.C., Jansse C.J., Barrell B., Turner C.M.R., Waters A.P., Sinden R.S.; "A comprehensive survey of the Plasmodium life cycle by genomic, transcriptomic, and proteomic analyses.";

RL Science 307:82-86(2005).

CC -!- CAUTION: The sequence shown here is derived from an EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.

CC EMBL; CA01000412; CAH94332.1; -; Genomic\_DNA.

DR Hypothetical protein.

KW NON\_TER

FT 1

SQ SEQUENCE 36 AA; 4144 MW; A481C901A7F72669 CRC64;

Query Match 95.2%; Score 20; DB 2; Length 36;

Best Local Similarity 100.0%; Pred. No. 4.3e+02; Mismatches 0; Indels 0; Gaps 0;

Query Match 95.2%; Score 20; DB 2; Length 36;

Q425P5 PLABE PRELIMINARY; PRT; 36 AA.

DT 13-SEP-2005 (TrEMBLrel. 31, Created)

DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)

DE Hypothetical protein.

GN ORENames=PB100986.00.0;

OS Plasmodium berghei.

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI\_TaxID=5821;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RA Hall N., Karras M., Raine J.D., Carlton J.M., Kooij T.W.A., Berriman M., Florens D., Janssen C.S., Pain A., Christophides G.K., James K., Rutherford K., Harris B., Harris H.E., Churcher C., Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J., Bidwell S.L., Rajandream M.A., Carucci D.J., Yates J.R., Kafatos F.C., Jansse C.J., Barrell B., Turner C.M.R., Waters A.P., Sinden R.S.; "A comprehensive survey of the Plasmodium life cycle by genomic, transcriptomic, and proteomic analyses.";

RL Science 307:82-86(2005).

CC -!- CAUTION: The sequence shown here is derived from an EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.

CC EMBL; CA01000412; CAH94332.1; -; Genomic\_DNA.

DR Hypothetical protein.

KW NON\_TER

FT 1

SQ SEQUENCE 36 AA; 4144 MW; A481C901A7F72669 CRC64;

Query Match 95.2%; Score 20; DB 2; Length 36;

Best Local Similarity 100.0%; Pred. No. 4.3e+02; Mismatches 0; Indels 0; Gaps 0;

Query Match 95.2%; Score 20; DB 2; Length 36;

Q425P5 PLABE PRELIMINARY; PRT; 36 AA.

DT 13-SEP-2005 (TrEMBLrel. 31, Created)

DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)

DE Hypothetical protein.

GN ORENames=PB100986.00.0;

OS Plasmodium berghei.

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI\_TaxID=5821;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RA Hall N., Karras M., Raine J.D., Carlton J.M., Kooij T.W.A., Berriman M., Florens D., Janssen C.S., Pain A., Christophides G.K., James K., Rutherford K., Harris B., Harris H.E., Churcher C., Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J., Bidwell S.L., Rajandream M.A., Carucci D.J., Yates J.R., Kafatos F.C., Jansse C.J., Barrell B., Turner C.M.R., Waters A.P., Sinden R.S.; "A comprehensive survey of the Plasmodium life cycle by genomic, transcriptomic, and proteomic analyses.";

RL Science 307:82-86(2005).

CC -!- CAUTION: The sequence shown here is derived from an EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.

CC EMBL; CA01000412; CAH94332.1; -; Genomic\_DNA.

DR Hypothetical protein.

KW NON\_TER

FT 1

SQ SEQUENCE 36 AA; 4144 MW; A481C901A7F72669 CRC64;

QY 1 RRLN 4  
DB 26 RRLN 29

Best Local Similarity 100.0%; Pred. No. 4.3e+02; Mismatches 0; Indels 0; Gaps 0;

Query Match 95.2%; Score 20; DB 2; Length 37;

Q425P5 PLABE PRELIMINARY; PRT; 37 AA.

DT 13-SEP-2005 (TrEMBLrel. 26, Created)

DT 13-SEP-2005 (TrEMBLrel. 26, Last sequence update)

DE Hypothetical protein.

GN ORENames=PB100986.00.0;

OS Plasmodium berghei.

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI\_TaxID=5821;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RA Hall N., Karras M., Raine J.D., Carlton J.M., Kooij T.W.A., Berriman M., Florens D., Janssen C.S., Pain A., Christophides G.K., James K., Rutherford K., Harris B., Harris H.E., Churcher C., Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J., Bidwell S.L., Rajandream M.A., Carucci D.J., Yates J.R., Kafatos F.C., Jansse C.J., Barrell B., Turner C.M.R., Waters A.P., Sinden R.S.; "A comprehensive survey of the Plasmodium life cycle by genomic, transcriptomic, and proteomic analyses.";

RL Science 307:82-86(2005).

CC -!- CAUTION: The sequence shown here is derived from an EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.

CC EMBL; CA01000412; CAH94332.1; -; Genomic\_DNA.

DR Hypothetical protein.

KW NON\_TER

FT 1

SQ SEQUENCE 37 AA; 4165 MW; D6339C3D80B98F60 CRC64;

Query Match 95.2%; Score 20; DB 2; Length 37;

Best Local Similarity 100.0%; Pred. No. 4.5e+02; Mismatches 0; Indels 0; Gaps 0;

Query Match 95.2%; Score 20; DB 2; Length 37;

Q425P5 PLABE PRELIMINARY; PRT; 37 AA.

DT 13-SEP-2005 (TrEMBLrel. 23, Created)

DT 13-SEP-2005 (TrEMBLrel. 23, Last sequence update)

DE Hypothetical protein.

GN ORENames=PB100986.00.0;

OS Plasmodium berghei.

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI\_TaxID=5821;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RA Hall N., Karras M., Raine J.D., Carlton J.M., Kooij T.W.A., Berriman M., Florens D., Janssen C.S., Pain A., Christophides G.K., James K., Rutherford K., Harris B., Harris H.E., Churcher C., Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J., Bidwell S.L., Rajandream M.A., Carucci D.J., Yates J.R., Kafatos F.C., Jansse C.J., Barrell B., Turner C.M.R., Waters A.P., Sinden R.S.; "A comprehensive survey of the Plasmodium life cycle by genomic, transcriptomic, and proteomic analyses.";

RL Science 307:82-86(2005).

CC -!- CAUTION: The sequence shown here is derived from an EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.

CC EMBL; CA01000412; CAH94332.1; -; Genomic\_DNA.

DR Hypothetical protein.

KW NON\_TER

FT 1

SQ SEQUENCE 37 AA; 4165 MW; D6339C3D80B98F60 CRC64;

Query Match 95.2%; Score 20; DB 2; Length 37;

Best Local Similarity 100.0%; Pred. No. 4.5e+02; Mismatches 0; Indels 0; Gaps 0;

Query Match 95.2%; Score 20; DB 2; Length 37;

Q425P5 PLABE PRELIMINARY; PRT; 37 AA.

DT 13-SEP-2005 (TrEMBLrel. 23, Created)

DT 13-SEP-2005 (TrEMBLrel. 23, Last sequence update)

DE Hypothetical protein.

GN ORENames=PB100986.00.0;

OS Plasmodium berghei.

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI\_TaxID=5821;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RA Hall N., Karras M., Raine J.D., Carlton J.M., Kooij T.W.A., Berriman M., Florens D., Janssen C.S., Pain A., Christophides G.K., James K., Rutherford K., Harris B., Harris H.E., Churcher C., Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J., Bidwell S.L., Rajandream M.A., Carucci D.J., Yates J.R., Kafatos F.C., Jansse C.J., Barrell B., Turner C.M.R., Waters A.P., Sinden R.S.; "A comprehensive survey of the Plasmodium life cycle by genomic, transcriptomic, and proteomic analyses.";

RL Science 307:82-86(2005).

CC -!- CAUTION: The sequence shown here is derived from an EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.

CC EMBL; CA01000412; CAH94332.1; -; Genomic\_DNA.

DR Hypothetical protein.

KW NON\_TER

FT 1

SQ SEQUENCE 37 AA; 4165 MW; D6339C3D80B98F60 CRC64;

RA	Jiang J.-X., Gu W.-Y., Zhang Y.-Q., Cai Z., Sheng H.-H., Yin H.-P.,	RT	"Evolutionary analysis of the well characterized endo16 promoter
RA	Zhang Y., Zhu G.-F., Wan M., Huang H.-L., Qian Z., Wang S.-Y., Ma W.,	RL	reveals substantial variation within functional sites.";
RA	Yao Z.-J., Shen Y., Qiang B.-Q., Xia Q.-C., Guo X.-K., Danchin A.,	DR	EMBL; DQ066806; AAY51735.1; -; Genomic_DNA.
RA	Saint Girons I., Somerville R.L., Wen Y.-M., Shi M.-H., Chen Z.,	DR	EMBL; DQ066809; AAY51738.1; -; Genomic_DNA.
RA	Xu J.-G., Zhao G.-P.,	DR	EMBL; DQ066811; AAY51740.1; -; Genomic_DNA.
RT	"Unique physiological and pathogenic features of Leptosira	DR	EMBL; DQ066812; AAY51741.1; -; Genomic_DNA.
RT	interrogans revealed by whole-genome sequencing.";	DR	EMBL; DQ066813; AAY51742.1; -; Genomic_DNA.
RL	Nature 422:888-893(2003).	DR	EMBL; DQ066816; AAY51745.1; -; Genomic_DNA.
DR	EMBL; AE011394; AAN49427.1; -; Genomic_DNA.	FT	EMBL; DQ066801; AAY51730.1; -; Genomic_DNA.
KW	Complete proteome.	FT	NON TER 39
SQ	SEQUENCE 37 AA; 4152 MW; CCL1BD48D35A495C CRC64;	SQ	SEQUENCE 39 AA; 4351 MW; 2AA9F34D78BE902A CRC64;
Query Match 95.2%; Score 20; DB 2; Length 37;			
Best Local Similarity 100.0%; Pred. No. 4.5e+02;			
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1 RRLN 4	QY	1 RRLN 4
Db		Db	
	5 RRLN 8		2 RRLN 5
RESULT 16			
Q4TTJ0_STRPU	PRELIMINARY;	Q4TTJ4_STRPU	PRELIMINARY;
ID	Q4TTJ0_STRPU	ID	Q4TTJ4_STRPU
AC	Q4TTJ0;	AC	Q4TTJ4;
DT	13-SEP-2005 (TrEMBLrel. 31, Created)	DT	13-SEP-2005 (TrEMBLrel. 31, Created)
DT	13-SEP-2005 (TrEMBLrel. 31, Last sequence update)	DT	13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT	13-SEP-2005 (TrEMBLrel. 31, Last annotation update)	DT	13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE	Extracellular protein (Fragment).	DE	Extracellular protein (Fragment).
GN	Name=endo16;	GN	Name=endo16;
OS	Strongylocentrotus purpuratus (Purple sea urchin).	OS	Strongylocentrotus purpuratus (Purple sea urchin).
OC	Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;	OC	Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC	Echinoidea; Euechinoidea; Echinacea; Echinoidea; Strongylocentrotidae;	OC	Echinoidea; Euechinoidea; Echinacea; Echinoidea; Strongylocentrotidae;
OC	Strongylocentrotus.	OC	Strongylocentrotus.
OX	NCBI_TaxID=7668;	OX	NCBI_TaxID=7668;
RN	[1]	RN	[1]
RP	NUCLEOTIDE SEQUENCE.	RP	NUCLEOTIDE SEQUENCE.
RX	PubMed=15937122;	RX	PubMed=15937122;
RA	Balhoff J.P., Wray G.A.;	RA	Balhoff J.P., Wray G.A.;
RT	"Evolutionary analysis of the well characterized endo16 promoter	RT	"Evolutionary analysis of the well characterized endo16 promoter
RT	reveals substantial variation within functional sites.";	RT	reveals substantial variation within functional sites.";
RL	Proc. Natl. Acad. Sci. U.S.A. 102:8591-8596(2005).	RL	Proc. Natl. Acad. Sci. U.S.A. 102:8591-8596(2005).
DR	EMBL; DQ066808; AAY51737.1; -; Genomic_DNA.	DR	EMBL; DQ066804; AAY51733.1; -; Genomic_DNA.
FT	NON TER 39	FT	NON TER 39
SQ	SEQUENCE 39 AA; 4429 MW; 2AA9F34D78AADA6F CRC64;	SQ	SEQUENCE 39 AA; 4337 MW; 7A9AF34D79FB9028 CRC64;
Query Match 95.2%; Score 20; DB 2; Length 39;			
Best Local Similarity 100.0%; Pred. No. 4.7e+02;			
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1 RRLN 4	QY	1 RRLN 4
Db		Db	
	2 RRLN 5		2 RRLN 5
RESULT 17			
Q4TTJ2_STRPU	PRELIMINARY;	Q4TTJ8_STRPU	PRELIMINARY;
ID	Q4TTJ2_STRPU	ID	Q4TTJ8_STRPU
AC	Q4TTJ2;	AC	Q4TTJ8;
DT	13-SEP-2005 (TrEMBLrel. 31, Created)	DT	13-SEP-2005 (TrEMBLrel. 31, Created)
DT	13-SEP-2005 (TrEMBLrel. 31, Last sequence update)	DT	13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT	13-SEP-2005 (TrEMBLrel. 31, Last annotation update)	DT	13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE	Extracellular protein (Fragment).	DE	Extracellular protein (Fragment).
GN	Name=endo16;	GN	Name=endo16;
OS	Strongylocentrotus purpuratus (Purple sea urchin).	OS	Strongylocentrotus purpuratus (Purple sea urchin).
OC	Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;	OC	Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC	Echinoidea; Euechinoidea; Echinacea; Echinoidea; Strongylocentrotidae;	OC	Echinoidea; Euechinoidea; Echinacea; Echinoidea; Strongylocentrotidae;
OC	Strongylocentrotus.	OC	Strongylocentrotus.
OX	NCBI_TaxID=7668;	OX	NCBI_TaxID=7668;
RN	[1]	RN	[1]
RP	NUCLEOTIDE SEQUENCE.	RP	NUCLEOTIDE SEQUENCE.
RX	PubMed=15937122;	RX	PubMed=15937122;
RA	Balhoff J.P., Wray G.A.;	RA	Balhoff J.P., Wray G.A.;

```
RA Balhoff J.P., Wray G.A.;
RT "Evolutionary analysis of the well characterized endol6 promoter
RT reveals substantial variation within functional sites.";
RL Proc. Natl. Acad. Sci. U.S.A. 102:8591-8596(2005).
DR EMBL; DQ066800; AAY51729.1; -; Genomic_DNA.
DR EMBL; DQ066802; AAY51731.1; -; Genomic_DNA.
DR EMBL; DQ066803; AAY51732.1; -; Genomic_DNA.
DR EMBL; DQ066810; AAY51739.1; -; Genomic_DNA.
DR EMBL; DQ066818; AAY51747.1; -; Genomic_DNA.
DR EMBL; DQ066799; AAY51728.1; -; Genomic_DNA.
FT NON_TER 39
SQ SEQUENCE 39 AA; 4365 MW; 2AA9F34D79FB902A CRC64;

Query Match 95.2%; Score 20; DB 2; Length 39;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4
Db 2 RRLN 5

RESULT 20
Q4TTK1_STRPU
ID Q4TTK1_STRPU PRELIMINARY; PRT; 39 AA.
AC Q4TTK1;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Extracellular protein (Fragment).
GN Name=endol6;
OS Strongylocentrotus purpuratus (Purple sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC Strongylocentrotus.
OC NCBI_TaxID=7668;
RN [1]_TaxID=7668;
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15937122;
RA Balhoff J.P., Wray G.A.;
RT "Evolutionary analysis of the well characterized endol6 promoter
RT reveals substantial variation within functional sites.";
RL Proc. Natl. Acad. Sci. U.S.A. 102:8591-8596(2005).
DR EMBL; DQ066797; AAY51726.1; -; Genomic_DNA.
FT NON_TER 39
SQ SEQUENCE 39 AA; 4399 MW; 2AA9F34D78BE8B2A CRC64;

Query Match 95.2%; Score 20; DB 2; Length 39;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4
Db 2 RRLN 5

RESULT 21
Q6546_9GAMA
ID Q6546_9GAMA PRELIMINARY; PRT; 39 AA.
AC Q6546;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE BBV B95-8 Cl(e) DNA with antigen coding ORF (Fragment).
OS Human herpesvirus 4 (Epstein-Barr virus).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gamaherpesvirinae; Lymphocryptovirus.
OC NCBI_TaxID=10376;
RN [1]_TaxID=10376;
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=88217505; PubMed=2835748;
RA Wallis D., Ferricaudet M., Gannon F.;
RT "The analysis of EBV proteins which are antigenic in vivo.";
```

```
RL Nucleic Acids Res. 16:2859-2872(1988).
DR EMBL; X07530; CAA30405.1; -; Genomic_DNA.
FT NON_TER 39
SQ SEQUENCE 39 AA; 4280 MW; 7937120F24AF5774 CRC64;

Query Match 95.2%; Score 20; DB 2; Length 39;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4
Db 10 RRLN 13

RESULT 22
Q4TTG6_STRDR
ID Q4TTG6_STRDR PRELIMINARY; PRT; 40 AA.
AC Q4TTG6;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Extracellular protein (Fragment).
GN Name=endol6;
OS Strongylocentrotus droebachiensis (Sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC Strongylocentrotus.
OC NCBI_TaxID=7671;
RN [1]_TaxID=7671;
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15937122;
RA Balhoff J.P., Wray G.A.;
RT "Evolutionary analysis of the well characterized endol6 promoter
RT reveals substantial variation within functional sites.";
RL Proc. Natl. Acad. Sci. U.S.A. 102:8591-8596(2005).
DR EMBL; DQ066862; AAY52171.1; -; Genomic_DNA.
FT NON_TER 40
SQ SEQUENCE 40 AA; 4540 MW; 70855D3D2101C533 CRC64;

Query Match 95.2%; Score 20; DB 2; Length 40;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4
Db 2 RRLN 5

RESULT 23
Q663L4_YERPS
ID Q663L4_YERPS PRELIMINARY; PRT; 40 AA.
AC Q663L4;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=pyV0052;
OS Yersinia pseudotuberculosis.
OG Plasmid pYV.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OC NCBI_TaxID=633;
RN [1]_TaxID=633;
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IP32953 / Serotype I;
RX PubMed=15358858; DOI=10.1073/pnas.04040120101;
RA Chain P.S.G., Carniel E., Larimer F.W., Lamerdin J., Stoutland P.O.,
RA Regala W.M., Georgescu A.M., Vergez L.M., Land M.L., Motin V.L.,
RA Brubaker R.R., Fowler J., Hinnebusch J., Marceau M., Medigue C.,
RA Simonet M., Chenal-Francois V., Souza B., Dacheux D., Elliott J.M.,
RA Derbise A., Hauser L.J., Garcia E.;
RT "Insights into the evolution of Yersinia pestis through whole-genome
RT comparison with Yersinia pseudotuberculosis.";
```

```

RL Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831(2004).
DR EMBL; BX936399; CAF25395.1.1; -; Genomic DNA.
KW Complete proteome; Hypothetical protein; Plasmid.
SQ SEQUENCE 40 AA; 4820 MW; 550F7DF129673PB0 CRC64;

Query Match 95.2%; Score 20; DB 2; Length 40;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4
Db |||||
25 RRLN 28

RESULT 24
Q7RL48 PLAYO
ID Q7RL48_PLAYO PRELIMINARY; PRT; 43 AA.
AC Q7RL48;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN Name=Plf02698;
OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=73239;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=L7XML;
RX MEDLINE=22255706; PubMed=12368865; DOI=10.1038/nature01099;
RA Carlton J.M., Angiuoli S.V., Suh B.B., Kooij T.W., Perte M.,
RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA Shallom S.J., van Aken S.E., Riedmuller S.B., Feldblyum T.V.,
RA Cho J.K., Quackenbush J., Sedegah M., Shoabi A., Cummings L.M.,
RA Florens L., Yates J.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
RA van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA Carucci D.J.;
RT "Genome sequence and comparative analysis of the model rodent malaria
RT parasite Plasmodium yoelii yoelii."
RL Nature 419:512-519(2002).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABL01000743; EAA22179.1.1; -; Genomic DNA.
KW Hypothetical protein.
SQ SEQUENCE 43 AA; 5154 MW; 42B52DAD652E255B CRC64;

Query Match 95.2%; Score 20; DB 2; Length 43;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4
Db |||||
15 RRLN 18

RESULT 25
Q8KEJ5 CHLTE
ID Q8KEJ5_CHLTE PRELIMINARY; PRT; 43 AA.
AC Q8KEJ5;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=CT0693;
OS Chlorobium tepidum.
OC Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
OC Chlorobaculum.
OX NCBI_TaxID=1097;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=TL5 / ATCC 49652 / DSM 12025;
RX MEDLINE=22103685; PubMed=12093901; DOI=10.1073/pnas.132181499;
RA Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,
RA Dodson R.J., DeBoy R.T., Gwinn M.L., Nelson W.C., Haft D.H.,
RA Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.F., Yang F.,
RA Holt I.E., Umayam L.A., Mason T.M., Brenner M., Shea T.P.,
RA Parksey D.S., Niernan W.C., Feldblyum T.V., Hansen C.L., Craven M.B.,
RA Radune D., Vamathevan J.C., Knouri H.M., White O., Gruber T.M.,
RA Ketchum K.A., Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.;
RT "The complete genome sequence of Chlorobium tepidum TLS, a
RT photosynthetic, anaerobic, green-sulfur bacterium."
RL Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).
DR EMBL; AB006470; AAM71931.1.1; -; Genomic DNA.
KW TIGR; CT0693; -;
RP Complete proteome; Hypothetical protein.
SQ SEQUENCE 43 AA; 5354 MW; EF0A76D438C1C2B7 CRC64;

Query Match 95.2%; Score 20; DB 2; Length 43;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4
Db |||||
8 RRLN 11

RESULT 26
Q5C5H2 SCHJA
ID Q5C5H2_SCHJA PRELIMINARY; PRT; 44 AA.
AC Q5C5H2;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein.
OS Schistosoma japonicum (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6182;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC Han Z.;
RA Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF809213; AAX25102.1.1; -; mRNA.
DR InterPro; IPR008919; Retrov_capsid_N.
KW Hypothetical protein.
SQ SEQUENCE 44 AA; 5441 MW; A96DEF5A34DAF9F1 CRC64;

Query Match 95.2%; Score 20; DB 2; Length 44;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4
Db |||||
34 RRLN 37

RESULT 27
Q72AJ8 DESVH
ID Q72AJ8_DESVH PRELIMINARY; PRT; 44 AA.
AC Q72AJ8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=DVU1994;
OS Desulfovibrio vulgaris (strain Hildenborough / ATCC 29579 / NCIMB
OS 8303).
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfovibrionales;
OC Desulfovibrionaceae; Desulfovibrio.
OX NCBI_TaxID=882;
RN [1]
RP NUCLEOTIDE SEQUENCE.

```



```

RX PubMed=15077118; DOI=10.1038/nbt959;
RA Heidelberg J.F., Seshadri R., Haveman S.A., Hemme C.L., Paulsen I.T.,
RA Kolonay J.F., Eissen J.A., Ward N.L., Methe B.A., Brinkac L.M.,
RA Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S., Madupu R.,
RA Nelson W.C., Sullivan S.A., Fouts D.E., Haft D.H., Selengut J.,
RA Peterson J.D., Davidsen T.M., Zafar N., Zhou L., Radune D.,
RA Dmitrov G., Hance M., Tran K., Khouri H.M., Gill J., Utterback T.R.,
RA Feldblyum T.V., Wall J.D., Voordouw G., Fraser C.M.;
RA "The genome sequence of the anaerobic, sulfate-reducing bacterium
RT Desulfovibrio vulgaris Hildenborough.";
RL Nat. Biotechnol. 22:554-559(2004).
DR ENBL; AE017314; AAG96470.1; -; Genomic_DNA.
DR TIGR; DVU1994; -;
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 44 AA; 5059 MW; EF626D0A31D46E3C CRC64;

Query Match 95.2%; Score 20; DB 2; Length 44;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4
DB 33 RRLN 36

RESULT 28
Q99IR3_9GEMI
ID Q99IR3_9GEMI PRELIMINARY; PRT; 47 AA.
AC Q99IR3;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Coat protein (Fragment).
OS Tomato leaf curl virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=28350;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Samretwanich K., Chiemsombat P., Kittipakorn K., Ikegami M.;
RT "Yellow leaf disease of muskmelon from Thailand caused by Tomato leaf
RT curl virus.";
RL Plant Dis. 84:707-707(2000).
DR ENBL; AB020976; BAB33262.1; -; Genomic_DNA.
KW Capsid protein.
FT NON_TER 47
SQ SEQUENCE 47 AA; 5281 MW; 60766BA44199E0C8 CRC64;

Query Match 95.2%; Score 20; DB 2; Length 47;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4
DB 19 RRLN 22

RESULT 29
Q99IR5_9GEMI
ID Q99IR5_9GEMI PRELIMINARY; PRT; 47 AA.
AC Q99IR5;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Coat protein (Fragment).
OS Tomato leaf curl virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=28350;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Samretwanich K., Chiemsombat P., Kittipakorn K., Ikegami M.;
RT "Tomato leaf curl geminivirus associated with cucumber yellow disease
RT in Thailand.";
RL J. Phytopathol. 148:615-617(2000).

```

```

RN NUCLEOTIDE SEQUENCE.
RP Samretwanich K., Chiemsombat P., Kittipakorn K., Ikegami M.;
RA "Yellow disease of cantaloupe and wax gourd from Thailand caused by
RT Tomato leaf curl virus.";
RL Plant Dis. 84:200-200(2000).
RN NUCLEOTIDE SEQUENCE.
RP Samretwanich K., Chiemsombat P., Kittipakorn K., Ikegami M.;
RA "Tomato leaf curl geminivirus associated with cantaloupe yellow leaf
RT disease in Thailand.";
RL World J. Microbiol. Biotechnol. 16:401-403(2000).
DR ENBL; AB017342; BAB33260.1; -; Genomic_DNA.
KW Capsid protein.
FT NON_TER 47
SQ SEQUENCE 47 AA; 5281 MW; 60766BA44199E0C8 CRC64;

Query Match 95.2%; Score 20; DB 2; Length 47;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4
DB 19 RRLN 22

RESULT 30
Q99IR1_9GEMI
ID Q99IR1_9GEMI PRELIMINARY; PRT; 48 AA.
AC Q99IR1;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE Coat protein (Fragment).
OS Soybean crinkle leaf virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=85753;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Samretwanich K., Kittipakorn K., Chiemsombat P., Ikegami M.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR ENBL; AB020977; BAB33264.1; -; Genomic_DNA.
DR GO; GO:0019028; F:structural molecule activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000650; Gem_coat_A1.
DR ProDom; PD000901; Gem_coat_A1; 1.
KW Capsid protein.
FT NON_TER 48
SQ SEQUENCE 48 AA; 5509 MW; 200765C58F2FDBA8 CRC64;

Query Match 95.2%; Score 20; DB 2; Length 48;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4
DB 19 RRLN 22

RESULT 31
PRTZ1_SCYCA
ID PRTZ1_SCYCA STANDARD; PRT; 50 AA.
AC P08433;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Protamine Z1 (Scyllorhinine Z1).
OS Scyllorhinus canicula (Spotted dogfish) (Spotted catshark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes;
OC Scyllorhinidae; Scyllorhinus.
OX NCBI_TaxID=7830;
RN [1]

```

```

Db          2 RRLN 5
|||||
RESULT 33
Q831H1 ENTRETF
ID Q831H1_ENTTFA PRELIMINARY; PRT; 50 AA.
AC Q831H1_
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=EF2538;
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
ON NCBI_TaxID=1351;
RX [1]
PC NUCLEOTIDE SEQUENCE.
RC STRAIN=V583 / ATCC 700802;
RX MEDLINE=22550857; PubMed=12663927; DOI=10.1126/science.1080613;
RA Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R.,
RA Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,
RA Tettelin H., Dodson R.J., Umayam L.A., Brinkac L.M., Beanan M.J.,
RA Daugherty S.C., DeBoy R.T., Durkin S.A., Kolonay J.P., Madupu R.,
RA Nelson W.C., Vamathevan J.J., Tran B., Upton J., Hansen T., Shetty J.,
RA Khouri H.M., Uterback T.R., Radune D., Ketchum K.A., Dougherty B.A.,
RA Fraser C.M.;
RT "Role of mobile DNA in the evolution of vancomycin-resistant
RT Enterococcus faecalis.";
RL Science 299:2071-2074(2003).
DR EMBL; AE016955; AAO82251.1; -; Genomic_DNA.
DR TIGR; EF2538; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 50 AA; 6083 MW; 9DB8BE43B9A11352 CRC64;

Query Match 95.2%; Score 20; DB 2; Length 50;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Qy 1 RRLN 4
|||||
Db 18 RRLN 21

RESULT 34
Q4T9A7 TETNG
ID Q4T9A7_TETNG PRELIMINARY; PRT; 50 AA.
AC Q4T9A7_
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 3 SCAP7569, whole genome shotgun sequence. (Fragment).
GN ORFNames=GSTNG00004841001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Daeulva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Blemont C., Skallat Z., Cattolico L., Poulet J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Winkler P., Lander E.S., Weissenbach J., Roest Crollius H.;

```

RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals  
 RL the early vertebrate proto-karyotype."; [2]  
 RN Nature 431:946-957(2004).  
 RP NUCLEOTIDE SEQUENCE.  
 RG Genoscope; Whitehead Institute Centre for Genome Research;  
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
 CC -1- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; CAAE01007608; CAF90525.1; -; Genomic\_DNA.  
 FT NON\_TER 1 1  
 FT NON\_TER 50 50  
 SQ SEQUENCE 50 AA; 6133 MW; F6C79EC09632BDC CRC64;  
 Query Match 95.2%; Score 20; DB 2; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 6.2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RRLN 4  
 Db 42 RRLN 45  
 RESULT 35  
 Q9ZSF6\_9SOLA PRELIMINARY; PRT; 51 AA.  
 AC Q9ZSF6;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Self-incompatibility ribonuclease S15 (Fragment).  
 OS Lycium andersonii.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 OC asterids; lamids; Solanales; Solanaceae; Lycium.  
 OX NCBI\_TaxID=87540;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Richman A.D., Kohn J.R.;  
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF105357; AD13085.1; -; mRNA.  
 DR GO; GO:0004521; F:endoribonuclease activity; IEA.  
 DR GO; GO:0016787; F:hydrolase activity; IEA.  
 DR GO; GO:0003723; F:RNA binding; IEA.  
 DR InterPro; IPR001568; RNase T2.  
 DR Pfam; PF00445; Ribonuclease\_T2; 1.  
 FT NON\_TER 1 1  
 FT NON\_TER 51 51  
 SQ SEQUENCE 51 AA; 6289 MW; CGA4248AC15ABA9E CRC64;  
 Query Match 95.2%; Score 20; DB 2; Length 51;  
 Best Local Similarity 100.0%; Pred. No. 6.3e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RRLN 4  
 Db 4 RRLN 7  
 RESULT 36  
 Q8RTK9\_LEUME PRELIMINARY; PRT; 51 AA.  
 AC Q8RTK9;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 OS Leuconostoc mesenteroides.  
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.  
 OX NCBI\_TaxID=1245;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.

RA Lee K.-H., Lee H.-J., Chang H.-C., Chung D.-K., Lee J.-H., Kim J.-H.;  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF420260; AAL77873.1; -; Genomic\_DNA.  
 KW Hypothetical protein.  
 SQ SEQUENCE 51 AA; 6177 MW; A929B57F6D8907D9 CRC64;  
 Query Match 95.2%; Score 20; DB 2; Length 51;  
 Best Local Similarity 100.0%; Pred. No. 6.3e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RRLN 4  
 Db 14 RRLN 17  
 RESULT 37  
 Q8U4Z9\_AGR75 PRELIMINARY; PRT; 51 AA.  
 AC Q8U4Z9;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE AGR\_L 771P.  
 GN OrderedLocusNames=AGR\_L\_771;  
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.  
 OX NCBI\_TaxID=176299;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=Cereon;  
 RX MEDLINE=21608551; PubMed=11743194; DOI=10.1126/science.1066803;  
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,  
 RA Quorllo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,  
 RA Hounell K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,  
 RA Wollam C., Allinger M., Doughty D., Scott C., Lapps C., Markelz B.,  
 RA Flanagan C., Crowell C., Gursion J., Lomo C., Sear C., Strub G.,  
 RA Cielo C., Slater S.;  
 RT "Genome sequence of the plant pathogen and biotechnology agent  
 RT Agrobacterium tumefaciens C58.";  
 RL Science 294:2323-2328(2001).  
 DR EMBL; AE008238; AAK88958.1; -; Genomic\_DNA.  
 DR FIR; D98179; D98179.  
 SQ SEQUENCE 51 AA; 5890 MW; 918E0B5D948C97DF CRC64;  
 Query Match 95.2%; Score 20; DB 2; Length 51;  
 Best Local Similarity 100.0%; Pred. No. 6.3e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RRLN 4  
 Db 34 RRLN 37  
 RESULT 38  
 Q57N53\_SALCH PRELIMINARY; PRT; 51 AA.  
 AC Q57N53;  
 DT 10-MAY-2005 (TrEMBLrel. 30, Created)  
 DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)  
 DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)  
 DE Putative inner membrane protein.  
 GN OrderedLocusNames=SCI952;  
 OS Salmonella cholerae-suis (Salmonella enterica).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Salmonella.  
 OX NCBI\_TaxID=591;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=SC-B67;  
 RX PubMed=15781495;  
 RA Chiu C.-H., Tang P., Chu C., Hu S., Bao Q., Yu J., Chou Y.-Y.,  
 RA Wang H.-S., Lee Y.-S.;

```

RT "The genome sequence of Salmonella enterica serovar Choleraesuis, a
RT highly invasive and resistant zoonotic pathogen.";
RL Nucleic Acids Res. 33:1690-1698(2005).
DR EMBL; AE017220; AAX65858.1; -; Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 51 AA; 6024 MW; 61270A2A7D28DD76 CRC64;

Query Match 95.2%; Score 20; DB 2; Length 51;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4
DB 33 RRLN 36

RESULT 39
Q8VKJ1 MYCTU
ID Q8VKJ1 MYCTU PRELIMINARY; PRT; 51 AA.
AC Q8VKJ1;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocNames=MT0576;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1773;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RX DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.B., Haft D.H.,
RA Hickey E.K., Kolonay J.F., Nelson W.C., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Uterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL J. Bacteriol. 184:5479-5490(2002).
RL EMBL; AE000516; AAK44799.1; -; Genomic_DNA.
DR TIGR; MT0576; -.
KW Hypothetical protein.
SQ SEQUENCE 51 AA; 5489 MW; B0803BCA77E3640B CRC64;

Query Match 95.2%; Score 20; DB 2; Length 51;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4
DB 2 RRLN 5

RESULT 40
QSP112 SALPA
ID QSP112 SALPA PRELIMINARY; PRT; 51 AA.
AC QSP112;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein.
GN OrderedLocNames=SFA0921;
OS Salmonella paratyphi-a.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=54388;
RN [1]
RP NUCLEOTIDE SEQUENCE.

```

```

RC STRAIN=ATCC 9150;
RX PubMed=15531882; DOI=10.1038/ng1470;
RA McClelland M., Sanderson K.E., Clifton S.W., Latreille P.,
RA Forwellik S., Sabo A., Meyer R., Bieri T., Ozersky P., McEllan M.,
RA Harkins C.R., Wang C., Nguyen C., Berghoff A., Elliott G.,
RA Kohlberg S., Strong C., Du F., Carter J., Kremizki C., Layman D.,
RA Leonard S., Sun H., Fulton L., Nash W., Miner T., Minx P.,
RA Delehaunt K., Fronick C., Magrini V., Nhan M., Warren W., Florea L.,
RA Spieth J., Wilson R.K.;
RA "Comparison of genome degradation in Paratyphi A and Typhi, human-
RT restricted serovars of Salmonella enterica that cause typhoid.";
RL Nat. Genet. 36:1268-1274(2004).
DR EMBL; CP000026; AAV76900.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 51 AA; 6011 MW; EC78D82A7D398933 CRC64;

Query Match 95.2%; Score 20; DB 2; Length 51;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4
DB 33 RRLN 36

RESULT 41
Q8Z5T2 SALT1
ID Q8Z5T2 SALT1 PRELIMINARY; PRT; 51 AA.
AC Q8Z5T2; Q7CAN6;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein STY2156.
GN OrderedLocNames=STY2156, t0928;
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608; DOI=10.1038/35101607;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Garra P., Parry C.,
RA Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrell B.G.;
RX "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Ty2 / AFCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RX DOI=10.1128/JB.185.7.2330-2337.2003;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodyanni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT and CT18.";
RT J. Bacteriol. 185:2330-2337(2003).
DR EMBL; AL627272; CAD05697.1; -; Genomic_DNA.
DR EMBL; AE016837; AAO68505.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 51 AA; 6011 MW; EC78D82A7D398933 CRC64;

Query Match 95.2%; Score 20; DB 2; Length 51;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4

```

Db 33 RRLN 36

# RESULT 42

Q82NT9 SALTY  
ID Q82NT9\_SALTY PRELIMINARY; PRT; 51 AA.  
AC Q82NT9;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Putative inner membrane protein.  
GN OrderedLocuNames-STMI948;  
OS Salmonella typhimurium.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Salmonella.  
OX NCBI\_TaxID=602;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=LT2;  
RX MEDLINE=21534948; PubMed=11677609; DOI=10.1038/35101614;  
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,  
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,  
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,  
RA Ryan B., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,  
RA Waterston R., Wilson R.K.;  
RA "Complete genome sequence of Salmonella enterica serovar Typhimurium  
RT LT2";  
RL Nature 413:852-856(2001).  
DR EMBL; AB008786; AAL20860.1; -; Genomic\_DNA.  
KW Complete proteome.  
SQ SEQUENCE 51 AA; 5995 MW; 8078D82A7D28CC27 CRC64;

Query Match 95.2%; Score 20; DB 2; Length 51;  
Best Local Similarity 100.0%; Pred. No. 6.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
Db 33 RRLN 36

# RESULT 43

Q96142 HUMAN  
ID Q96142\_HUMAN PRELIMINARY; PRT; 52 AA.  
AC Q96142;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Hypothetical protein (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Brain;  
RA Strausberg R.;  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC007841; AH07841.1; -; mRNA.  
KW Hypothetical protein.  
FT NON\_TER 1  
SQ SEQUENCE 52 AA; 5746 MW; 13C8F9B7A0532365 CRC64;

Query Match 95.2%; Score 20; DB 2; Length 52;  
Best Local Similarity 100.0%; Pred. No. 6.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
Db 38 RRLN 41

# RESULT 44

Q54BT9 DICDI  
ID Q54BT9\_DICDI PRELIMINARY; PRT; 52 AA.  
AC Q54BT9;  
DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
DE Hypothetical protein.  
GN ORFNames=DOB0191934;  
OS Dictyostelium discoideum (Slime mold).  
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.  
OX NCBI\_TaxID=44689;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=AX4;  
RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,  
RA Sugang R., Berriman M., Song J., Olsen R., Szafranski K., Xu Q.,  
RA Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,  
RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,  
RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,  
RA Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Bason N.,  
RA Farbrother P., Desany B., Just E., Morio T., Rost R., Churcher C.,  
RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,  
RA Muzny D., Mourier T., Pain A., Lu M., Harper D., Lindsay R.,  
RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,  
RA Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,  
RA Loulseghe H., Mungall K., Oliver K., Price C., Quail M.A., Sanders M.,  
RA Urushihara H., Hernandez J., Rabinowitsch E., Steffen D., Tivey A.,  
RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,  
RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,  
RA Shaulsky G., Schleicher M., Weinstock G., Rosenthal A., Cox E.C.,  
RA Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.R.,  
RA Williams J., Dear P.H., Noegel A.A., Barrell B., Kuspa A.;  
RT "The genome of the social amoeba Dictyostelium discoideum";  
RL Nature 0:0-0(2005).  
CC -! CAUTION: The sequence shown here is derived from an  
CC preliminary data.  
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
DR EMBL; AAFI01000284; EAL60727.1; -; Genomic\_DNA.  
KW Hypothetical protein.  
SQ SEQUENCE 52 AA; 6226 MW; 1C226192CE0CE606 CRC64;

Query Match 95.2%; Score 20; DB 2; Length 52;  
Best Local Similarity 100.0%; Pred. No. 6.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
Db 19 RRLN 22

# RESULT 45

Q9ZSG7\_9SOLA  
ID Q9ZSG7\_9SOLA PRELIMINARY; PRT; 53 AA.  
AC Q9ZSG7;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Self-incompatibility ribonuclease S4 (Fragment).  
OS Lycium andersonii.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
OC asterids; lamiids; Solanales; Solanaceae; Lycium.  
OX NCBI\_TaxID=87540;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Richman A.D., Kohn J.R.;  
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF105346; AAD13074.1; -; mRNA.  
DR GO; GO:0004521; F:endoribonuclease activity; IEA.  
DR GO; GO:0016787; F:hydrolase activity; IEA.  
DR GO; GO:0003723; F:RNA binding; IEA.

```
DR InterPro: IPR001568; RNase_T2.
DR Pfam: PF00445; Ribonuclease_T2; 1.
FT NON_TER 1
FT NON_TER 53
SQ SEQUENCE 53 AA; 6470 MW; 9ACDEBF804A414E CRC64;

Query Match 95.2%; Score 20; DB 2; Length 53;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4
Db 4 RRLN 7

RESULT 46
Q8EDQ1_SHEON
ID Q8EDQ1_SHEON PRELIMINARY; PRT; 54 AA.
AC Q8EDQ1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein SO2692.
GN OrderedLocusNames=SO2692;
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Shewanellaceae; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=NR-1;
RX MEDLINE=22297686; PubMed=12368813; DOI=10.1038/nbt749;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward M.L., Methe B.A.,
RA Clayton R.A., Meyer T., Tsapin A., Scott J., Beanan M.J.,
RA Brinkac L.M., Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S.,
RA Haft D.H., Kolonay J.P., Madupu R., Peterson J.D., Unayam L.A.,
RA White O., Wolf A.M., Vamathevan J.J., Weidman J.F., Imprial M.,
RA Lee K., Berry K.J., Lee C., Mueller J., Khouri H.M., Gill J.,
RA Utterback T.R., McDonald L.A., Feldblyum T.V., Smith H.O.,
RA Venter J.C., Nealon K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis."
RL Nat. Biotechnol. 20:1118-1123 (2002).
DR EMBL; AE015709; AAN55720.1; -; Genomic_DNA.
DR TIGR; SO2692; -.
KW Complete proteome.
SQ SEQUENCE 54 AA; 5964 MW; 8AFC3DAB6DBA2D37 CRC64;

Query Match 95.2%; Score 20; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4
Db 37 RRLN 40

RESULT 47
Q8D157_SYNEL
ID Q8D157_SYNEL PRELIMINARY; PRT; 54 AA.
AC Q8D157;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Tsr1734 protein.
GN OrderedLocusNames=tsr1734;
OS Synechococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=32046;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BP-1;

RX MEDLINE=2225144; PubMed=12240834;
RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA Watanabe A., Iriguchi M., Kawashima K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
RA Shimo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the thermophilic cyanobacterium
RT Thermosynechococcus elongatus BP-1."
RL DNA Res. 9:123-130 (2002).
DR EMBL; BA000039; BAC09286.1; -; Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 54 AA; 6487 MW; 2F9494243FB80E67 CRC64;

Query Match 95.2%; Score 20; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4
Db 32 RRLN 35

RESULT 48
Q9E3W1_TYLCV
ID Q9E3W1_TYLCV PRELIMINARY; PRT; 54 AA.
AC Q9E3W1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 20, Last annotation update)
DE Coat protein (Fragment).
OS Tomato yellow leaf curl virus (TYLCV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10832;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Ying Z., Davis M.J.;
RT "Partial characterization and host range of tomato yellow leaf curl
RT virus in south Florida."
RL Annu. Meet. Fla. State Hort. Soc. 0:0-0 (2000).
DR EMBL; AF260333; AAG23626.1; -; Genomic_DNA.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000650; G:coat AR1.
DR ProDom; PD000901; G:coat_AR1; 1.
KW Capsid protein.
FT NON_TER 54
SQ SEQUENCE 54 AA; 6349 MW; A626B63B92E07376 CRC64;

Query Match 95.2%; Score 20; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4
Db 19 RRLN 22

RESULT 49
Q4GXV3_9GEMI
ID Q4GXV3_9GEMI PRELIMINARY; PRT; 54 AA.
AC Q4GXV3;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Coat protein (Fragment).
GN Name=AV1;
OS Tomato geminivirus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus;
OC unclassified Begomovirus.
OX NCBI_TaxID=32610;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Li Z., Zhou X.;
RT "Molecular characterization of tomato-infecting gemoviruses in Yunnan,
```

RT China.";  
 RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AM048873; CAJ15534.1; -; Genomic\_DNA.  
 KW Capsid protein.  
 FT NON TER 54  
 SQ SEQUENCE 54 AA; 6249 MW; 23B2A35D6D8759F0 CRC64;

Query Match 95.2%; Score 20; DB 2; Length 54;  
 Best Local Similarity 100.0%; Pred. No. 6.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
 ||||  
 Db 19 RRLN 22

RESULT 50

Q4GXV7\_9GEMI  
 ID Q4GXV7\_9GEMI PRELIMINARY; PRT; 55 AA.  
 AC Q4GXV7;  
 DT 13-SEP-2005 (TREMBLrel. 31, Created)  
 DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)  
 DT 13-SEP-2005 (TREMBLrel. 31, Last annotation update)  
 DE Coat protein (Fragment).  
 GN Name=AV1;  
 OS Tomato geminivirus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus;  
 OC unclassified Begomovirus.  
 OX NCBI\_TaxID=32610;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Li Z., Zhou X.;  
 RT "Molecular characterization of tomato-infecting gemoviruses in Yunnan, China."  
 RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AM048871; CAJ15530.1; -; Genomic\_DNA.  
 KW Capsid protein.  
 FT NON TER 55  
 SQ SEQUENCE 55 AA; 6324 MW; 70A2D5A34EF03191 CRC64;

Query Match 95.2%; Score 20; DB 2; Length 55;  
 Best Local Similarity 100.0%; Pred. No. 6.9e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRLN 4  
 ||||  
 Db 19 RRLN 22

Search completed: January 25, 2006, 18:41:54  
 Job time : 103 secs

**THIS PAGE BLANK (USPTO)**